

**UNIVERSIDAD COMPLUTENSE DE MADRID**

**FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES**



**TESIS DOCTORAL**

**Modelo para determinar la cuota en tienda de un producto en el punto  
de venta**

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

**Tomás Curto González**

Director

Enrique García Pérez

**Madrid, 2016**



UNIVERSIDAD COMPLUTENSE DE MADRID

# Modelo para determinar la cuota en tienda de un producto en el punto de venta

---

Tesis doctoral

**Autor: Tomás Curto González**

**Director: Enrique García Pérez**

**[Madrid, 28 de septiembre de 2015]**

*A Dalmacia,  
por hacer realidad lo imposible*

## ÍNDICE.

<b>0. RESUMEN Y ABSTRACT .....</b>	<b>7</b>
<b>1. INTRODUCCIÓN .....</b>	<b>14</b>
<b>2. MODELOS DE MARKETING .....</b>	<b>51</b>
2.1. Tipologías de Modelos de marketing .....	57
2.1.1. Modelos de Toma de Decisiones de los Consumidores .....	57
2.1.2. Modelos de Marketing Mix.....	58
2.1.3. Modelos de Marketing Centrados en el Cliente .....	59
2.1.4. Otras aproximaciones.....	60
2.2. Modelos Relacionados con el Modelo que se presenta en esta Tesis.....	60
2.2.1. Modelos de Publicidad .....	61
2.2.2. Modelos de Promoción de Ventas.....	64
2.2.3. Modelos de Gestión de Ventas.....	68
2.2.4. Modelos de Respuesta Competitiva .....	69
2.3. ¿Qué tiene en común y en qué se diferencia el modelo que se presenta en esta tesis con modelos analizados? .....	79
<b>3. ¿POR QUÉ DATOS DE PANEL?.....</b>	<b>84</b>
3.1. El Modelo Lineal de Datos de Panel .....	85
3.1.1. Método MCO Agrupado.....	86
3.1.1.1. Hipótesis para estimar $\beta$ de forma consistente por el Método MCO Agrupado .....	87
3.1.1.2. Matriz de Varianzas Robusta Asintóticamente.....	89
3.1.1.3. Test de Autocorrelación y Heteroscedasticidad después de estimar por MCO Agrupados .....	90
3.1.2. Modelos de Datos de Panel con Efectos no Observados.....	92
3.1.3. El problema de las Variables Omitidas.....	93
3.1.4. Efectos Fijos o Aleatorios .....	95
3.1.5. Estimación de Modelos de Efectos no Observados con MCO Agrupados .....	96
3.1.6. Método de Efectos Aleatorios.....	97
3.1.6.1. Estimador Robusto de la Matriz de Varianzas .....	104
3.1.6.2. Análisis General de Mínimos Cuadrados Generalizados Factible (MCGF) .....	106

3.1.7. Métodos de Efectos Fijos .....	107
3.1.7.1. Regresión de la Variable Ficticia.....	116
3.1.7.2. Autocorrelación y Estimador Robusto de la Matriz de Varianzas.....	118
3.1.7.3. Efectos Fijos con Mínimos Cuadrados Generalizados (MCG) ...	121
3.1.8. Método de Primeras Diferencias .....	123
3.1.8.1. Inferencia .....	123
3.1.8.2. Matriz De Varianza Robusta .....	127
3.1.8.3. Contraste De Autocorrelación .....	128
3.1.9. Comparación de Estimadores .....	128
3.1.9.1. Efectos Fijos versus Primeras Diferencias.....	128
3.1.9.2. Relación entre los Estimadores de Efectos Aleatorios y Efectos Fijos .....	131
3.1.9.3. Test de Hausman para comparar Estimadores de Efectos Aleatorios y Efectos Fijos.....	136
<b>4. CONTRASTES EN DATOS DE PANEL .....</b>	<b>141</b>
4.1. Heteroscedasticidad y Autocorrelación en Datos de Panel.....	142
4.1.1. Heteroscedasticidad .....	142
4.1.1.1. Detección de la Heteroscedasticidad .....	142
4.1.1.1.1. Contraste de Heteroscedasticidad basado en el Test de Bartlett .....	144
4.1.1.1.2. Contraste de Heteroscedasticidad basado en el Test de Levene.....	146
4.1.1.1.3. Corrección de la Heteroscedasticidad .....	147
4.1.2. Autocorrelación.....	150
4.1.2.1. Detección de la Autocorrelación .....	150
4.1.2.1.1. Test LM Conjunto para Autocorrelación y Efectos Individuales Aleatorios.....	151
4.1.2.1.2. Test LM Conjunto para Autocorrelación de Primer Orden en un Modelo de Efectos Aleatorios .....	153
4.1.2.1.3. Test LM Conjunto para Autocorrelación de Primer Orden en un Modelo De Efectos Fijos.....	156
4.1.2.1.4. Test Durbin – Watson para Datos de Panel .....	158
4.1.2.1.5. Test para diferenciar un AR(1) frente un MA(1) en un Modelo de Componentes del Error.....	160

4.1.2.1.6. Test Alternativo tipo BGT Para Contrastar un AR(1) frente Un MA(1) .....	162
4.1.2.2. Corrección de la Autocorrelación .....	169
4.1.2.2.1. Autocorrelación generada por un proceso AR(1) en un Modelo de Efectos Aleatorios .....	169
4.1.2.2.2. Autocorrelación generada por un proceso AR(1) en un Modelo de Efectos Fijos .....	178
4.2. Un Planteamiento Alternativo.....	182
4.3. Test de Especificación de Hausman .....	186
4.4. Test de Agrupabilidad .....	190
4.5. Contraste Reset (Regression Specification Error Test) o de Error de Especificación .....	194
<b>5. ESPECIFICACIÓN DEL MODELO .....</b>	<b>197</b>
5.1. Variable a explicar .....	197
5.2. Factores explicativos .....	201
5.3. Variables explicativas .....	203
5.4. ESTIMACIÓN DEL MODELO .....	211
5.4.1. Estimación estática del modelo .....	212
5.4.1.1. Estadísticos descriptivos.....	212
5.4.1.2. Correlaciones entre las variables.....	214
5.4.1.3. Variabilidad del modelo.....	216
5.4.2. ESTIMACIÓN POR MCO AGRUPADOS .....	220
5.4.3. MODELO DE EFECTOS INDIVIDUALES ESPECÍFICOS .....	227
5.4.3.1. Estimador Within o de Efectos Fijos (EF).....	228
5.4.3.2. Estimador de Efectos Aleatorios (EA) .....	230
5.4.3.3. Estimación Dinámica del modelo .....	239
5.4.3.3.1. Conclusiones de este modelo.....	243
5.4.4. Modelo de datos de panel por establecimientos o por formatos.....	245
5.4.4.1. Contraste de Agrupabilidad por Establecimientos .....	246
5.4.4.2. Contraste de Agrupabilidad por Formatos .....	249
<b>6. ¿QUÉ VARIABLES INFLUYEN EN LOS MODELOS? .....</b>	<b>252</b>
6.1. Estimación del Modelo Dinámico de Efectos Individuales Específicos Para El Formato hasta 350 g.....	253
6.1.1. Variables explicativas .....	254

6.2. Estimación del Modelo Dinámico de Efectos Individuales Específicos Formato de 350 a 499 g.....	268
6.3. Estimación del Modelo Dinámico de Efectos Individuales Específicos Formato de 500 A 899 g. ....	280
6.4. Estimación del Modelo Dinámico de Efectos Individuales Específicos Formato de 900 A 1199 g. ....	289
6.5. Estimación del Modelo Dinámico de Efectos Individuales Específicos Formato de 1200 A 1499 g. ....	301
6.6. Estimación del Modelo Dinámico de Efectos Individuales Específicos Tipo de Establecimiento Hipermercados. ....	315
<b>7. CONCLUSIONES .....</b>	<b>328</b>
7.1. ¿Qué Variables resultan más Influyentes en los Distintos Modelos?.....	335
7.1.1. Elasticidad en el modelo Hasta 350 g. ....	336
7.1.2. Elasticidad en el modelo de 350 A 499 g. ....	341
7.1.3. Elasticidad en el modelo de 500 A 899 g. ....	346
7.1.4. Elasticidad en el modelo de 900 A 1199 g. ....	349
7.1.5. Elasticidad en el modelo de 1200 A 1499 g.....	356
7.1.6. Elasticidad en el modelo de Hipermercados.....	361
<b>8. BIBLIOGRAFÍA .....</b>	<b>368</b>

## **0.- RESUMEN**

### **Título: Modelo para determinar la cuota en tienda de un producto en el punto de venta**

La elaboración de modelos de marketing ha sido una labor emprendida por distintos autores que comienza en la década de los 60 y llega hasta la actualidad. Sin embargo la mayor parte de estos modelos, han pasado desapercibidos para los profesionales del marketing.

El objetivo fundamental de esta tesis es construir un modelo de marketing que determine que variables, propias y de la competencia, son las que influyen en la cuota en tienda de una marca.

Para lograr el objetivo, se ha comenzado, en primer lugar, por revisar las aportaciones realizadas en el área de los modelos de marketing. En segundo lugar se ha realizado una clasificación de los modelos, para ubicar mejor al modelo objeto de esta tesis.

La metodología para construir el modelo ha venido determinada por el tipo de dato disponible. Los datos utilizados han sido proporcionados por Nielsen. Son datos agregados que muestran la evolución en el tiempo de diferentes magnitudes de marketing. Esta agregación puede realizarse por formato y/o por tipo de establecimiento.

La primera cuestión a contestar es: ¿Existe un modelo para cada establecimiento/formato o por el contrario un modelo único y global? Un modelo único, supone aceptar una realidad comercial idéntica independientemente del formato y/o el establecimiento; una realidad demasiado simple. La alternativa es admitir un modelo diferente para cada formato y establecimiento; una realidad compleja, que implica una mayor dificultad para analizar la realidad comercial.

Pero ¿existe algún punto intermedio entre ambas hipótesis?, ¿existe alguna técnica que permita encontrar puntos en común entre formatos y establecimientos y haga



posible formular modelos más adaptados a las diferentes realidades comerciales pero sin estar sometidos a la necesidad de un modelo para cada establecimiento?

Es aquí donde los modelos de datos de panel se han mostrado útiles, mostrando evidencias en contra a la existencia de un único modelo y proporcionando una guía para agrupar los modelos. De manera que sin tener un modelo para cada establecimiento y formato si ha podido encontrarse un punto en común.

La variable a estudiar es la cuota en tienda, la cuota de mercado en los establecimientos donde está presente la marca. Son datos del mercado español de cacao soluble. Nielsen ha borrado cualquier referencia a las marcas que operan en el mercado para garantizar el anonimato de sus clientes.

¿Por qué razón la cuota en tienda? Cuando se trabaja con datos Nielsen, la evolución de un mercado viene dada por:

- La distribución, número de establecimientos que trabajan el producto y
- La cuota en tienda o cuota de mercado en los puntos de venta donde está presente la marca.

Trabajar con la cuota de mercado añade otra fuente de heterogeneidad a los datos: la distribución. Por esta razón es preferible trabajar con una variable más homogénea.

La cuota en tienda de la Marca 1 es la variable a explicar. Las variables explicativas utilizadas están relacionadas con el entorno competitivo de la Marca 1 formado por el PRINCIPAL COMPETIDOR, RESTO DE FABRICANTES Y MARCAS DEL DISTRIBUIDOR. Dentro de este entorno se definen las variables explicativas, referidas al precio, la distribución y las promociones.

Desde la óptica del marketing, esta conclusión, supone admitir que NO hay un único modelo de gestión para todos los formatos y establecimientos. Y si es así, ¿cuántos modelos existen? Para dar respuesta a la pregunta se dividió la muestra en función

de dos criterios: los establecimientos y los formatos. Se consideraron dos unidades de agrupación diferentes:

- Los establecimientos: Hipermercados, Supermercados de 1000 a 2499 m, Supermercados de 400 a 999 m y Supermercados de 100 a 399 m.
- Los formatos: Hasta 350 g, De 350 a 499 g, De 500 a 899 g, De 900 a 1200 g y De 3000 a 4999 g.

Los resultados de la investigación respecto a los establecimientos indican que:

- En los Supermercados pequeños y en los hipermercados se acepta que existe un único modelo para los formatos comercializados.
- En los Supermercados medianos y grandes se rechaza la hipótesis nula; no hay un modelo único para todos, sino un modelo para cada formato.

Cuando se repite el análisis, pero por formatos, los resultados indican que:

- Hay 5 modelos distintos, uno para cada formato. Estos modelos son válidos para todos los establecimientos.
- El único formato donde se rechaza la hipótesis es en el de 3000 a 4999 g.

Ahora podemos afirmar que el comportamiento de la empresa refleja dos dimensiones que se superponen. Una dimensión vertical, que evidencia un tratamiento homogéneo por formato independientemente del establecimiento y una dimensión horizontal con un tratamiento homogéneo por establecimiento y no por formato.

En conclusiones, podemos afirmar que:

Es posible construir diferentes modelos que expliquen el comportamiento comercial de las marcas en el punto de venta. Modelos que determinan qué variables influyen y el peso de esta influencia.

No existe una política comercial única para la marca, sino que la política comercial es distinta según los formatos y los clientes. Y por último, la elasticidad de las variables significativas, toma valores distintos entre ellas, pero además difieren de un tipo de establecimiento a otro.

## **ABSTRACT**

### **Title: Model to determine the fee in store for a product at shop**

Marketing modeling has been undertaken by different authors since the 1960s, until the present. However, most of these models have gone unnoticed for marketing professionals.

The main objective of this dissertation is to build a marketing model that will determine which variables have an impact in brand share within different types of stores.

To address this goal, two tasks have been undertaken: first, literature in the area of marketing models has been revised. Second, a classification of models has been attempted, so that the model proposed by the dissertation can be placed.

The methodology to build the model has been determined by the type of data available. The data used were provided by Nielsen. They are aggregated data showing the evolution over time of different magnitudes of marketing. This aggregation can be done by format and/or by type of establishment.

The dissertation aims to answer one question raised: Is there a model for each establishment / format or, on the contrary, only a single global model? On the one hand, the existence of only one model supposes the acceptance of a homogeneous and unique commercial reality, regardless of format and/or establishment. Therefore, the “just one model” hypothesis involves the notion of too a simple reality. On the other, the alternative hypothesis means the existence of different models for different formats and establishments. This alternative possibility entails a more complex reality as a framework, and it requires a more complex and detailed analysis of day to day commercial realities/practices.

However, could there be a third way between the two mentioned hypotheses? Is it possible to identify a technique applicable to different formats and establishments? Such a technique will identify a common ground among formats and establishments,

thus allowing us the tailoring of models to different business realities while freeing us from the need to build a different one for each establishment.

This is where the panel data models have been very useful, showing contrary evidence regarding the existence of a single model and providing guidance to group models. Therefore, some common ground has been found, without the need for a model and format for each establishment. We have decided to consider as variable the “in-store market share” vs. market share variable. The product chosen is a Spanish cocoa powder, so the data analyzed come from there. Nielsen has deleted any reference to brands operating in the market to ensure the anonymity of their clients.

Why choosing “in-store market share” as variable? When working with Nielsen data, the development of a market is measured by:

- Distribution, that is, the number of establishments selling the product and
- “in-store market share”, or sales of the product among retail outlets where the brand is present.

Nevertheless, “in-store market share” option adds yet another source of heterogeneity to the data: distribution. That is why working with a more homogeneous variable is preferable.

Brand 1’s “In-store market share is the dependent variable. The explanatory variables are related to Brand 1 competitive environment, which consists of the MAIN COMPETITOR, OTHER MANUFACTURERS and STORE BRANDS. Within this competitive context, price, distribution and promotions are defined.

From a marketing point of view, the study concludes that there is not a single valid model for all sizes and establishments. Therefore, the goal is to find out how many models would be required. In order to achieve fulfill it, the sample was divided by two criteria: establishments and product sizes. Two different groups were found:

1. Establishments: hypermarkets, supermarkets (1000 to 2499 square ft.)  
Supermarkets (400 to 999 Square ft.), and supermarkets (100 to 399 square ft.).
2. Sizes: Up to 350 gr.; 350 to 499 gr.; 500 to 899 gr., 900 to 1200 gr, and 3000 to 4999 gr.

Research conducted on group 1 showed that small supermarkets (100 to 399 square ft.) and hypermarkets require only one model to explain product behavior. Medium and big size supermarkets require more than one model to explain product behavior.

Research conducted on group 2 indicated that there are 5 different models: one per size. These models are valid in all types of establishment, except in regard to 3000 to 4999 format, in which case the hypothesis is rejected.

Research, thus, proves that brand behavior encompasses two overlapping dimensions. A vertical one, which shows a homogenous conduct within the size, regardless of the type of establishment; and a horizontal one, which implies a homogeneous behavior within the establishment.

The study offers as final conclusion the possibility of building different models to explain behavior in establishment, taking into account explanatory variables and the weight of them. There is no just one product behavior within the brand. On the contrary, this product behavior is different depending on sizes and clients. Elasticity of variables changes, depending on type of establishment.

## **1.- INTRODUCCIÓN**

El objetivo fundamental de esta tesis es la construcción de un modelo de marketing que permita determinar que variables, tanto propias como de la competencia, son las que influyen en la cuota de mercado de un producto. Y de esta manera poder cumplir un objetivo fundamental en marketing: controlar a la competencia y reaccionar ante posibles acciones de ella.

El proceso hasta la construcción del modelo se ha convertido en un largo viaje vital, en un primer momento, en la génesis de la idea, la idea respecto a la tesis te parece absolutamente novedosa y que a nadie ha podido ocurrírsele algo parecido. En poco tiempo sales de la ignorancia y te das cuenta que tu idea es una más de las muchas que ya existen, por lo tanto es el momento de reconsiderarla y reubicarla, y determinar que aportación puede realizar tu trabajo respecto a lo que ya existe. Por lo tanto, y como primer paso, es necesario conocer que aportaciones se han realizado en tu campo de estudio.

El capítulo 1 cubre precisamente esta necesidad. En él se pasa revista a las aportaciones realizadas en el área de los modelos de marketing. Es un capítulo transcendental, quizás por lo poco conocido que resulta esta disciplina dentro de los profesionales del marketing y de las escuelas de negocios. El capítulo comienza con una evolución histórica de los modelos de marketing, desde los años 60 con los modelos de Bass (1961), Frank (1962) y Buzzel (1964) hasta las últimas aportaciones realizadas en la primera década del siglo XXI. Posteriormente a esta introducción histórica de los modelos de marketing, se realiza su clasificación. Esta clasificación es importante porque permite ubicar el modelo que se presenta en esta tesis.

Utilizando la clasificación propuesta por Wierenga (2008:10), los tipos de modelos de marketing pueden clasificarse como:

- Modelos de Toma de Decisiones de los Consumidores. Se trata de modelos que intentan explicar cómo los consumidores toman las decisiones de compra.
- Modelos de Marketing Mix. Modelos que explican cómo influyen los instrumentos de marketing en las ventas. El tema es interesante por dos motivos. El primero, su difusión; es el campo de utilización más importante de los modelos de marketing desde su inicio en los años 70. El segundo está relacionado con la tesis, ya que propone un modelo donde las variables explicativas son elementos del marketing mix. Dentro de estos modelos, Wierenga (2008: 11) los clasifica según el instrumento de marketing utilizado para explicar las ventas, de esta forma habla de los siguientes tipos de modelos:
  - Modelos de Publicidad, son modelos que explican cómo influye la publicidad en las ventas de la marca.
  - Modelos de Promoción de Ventas. En este caso lo interesante es modelizar y medir los efectos de la promoción de ventas.
  - Modelos para Decisiones de Gestión de Ventas. Se intenta explicar cómo las ventas responden ante cambios en la localización de recursos de ventas, en el diseño del territorio o en la estructura de la fuerza de ventas.
  - Modelos de Respuesta Competitiva. Se trata de explicar la respuesta de las ventas a las variables de marketing mix de la competencia.
- Modelos de Marketing Centrados en el Cliente. Su desarrollo se ha debido a la influencia creciente del cliente en el marketing unido a la gran cantidad de información disponible. El interés recae en procesar toda esa información e intentar explicar el comportamiento de cada uno de los clientes. Dentro de este grupo de modelos podemos encontrar:
  - Los modelos del Valor del Cliente.
  - Los modelos de Decisión para la Gestión de Relaciones con Clientes (CRM).



- Otras aproximaciones. Aquí se encuentran aportaciones con un desarrollo limitado en el área de marketing, como:
  - Modelos de Series Temporales en Marketing.
  - Redes Neuronales y Algoritmos Genéticos en Marketing.
  - Modelos desarrollados para algún sector específico (Cine, Industria Farmacéutica ...)
  - Modelos donde se analizan los efectos financieros del marketing en las empresas....

El modelo desarrollado en esta tesis puede inscribirse dentro de los modelos de marketing mix; aun así se realiza una aproximación más cercana a los modelos de publicidad, de promoción de ventas, de gestión de ventas, de respuesta competitiva y de series temporales en marketing para analizar las semejanzas y diferencias que el modelo propuesto en la tesis tiene con ellos así como la aportación que supone.

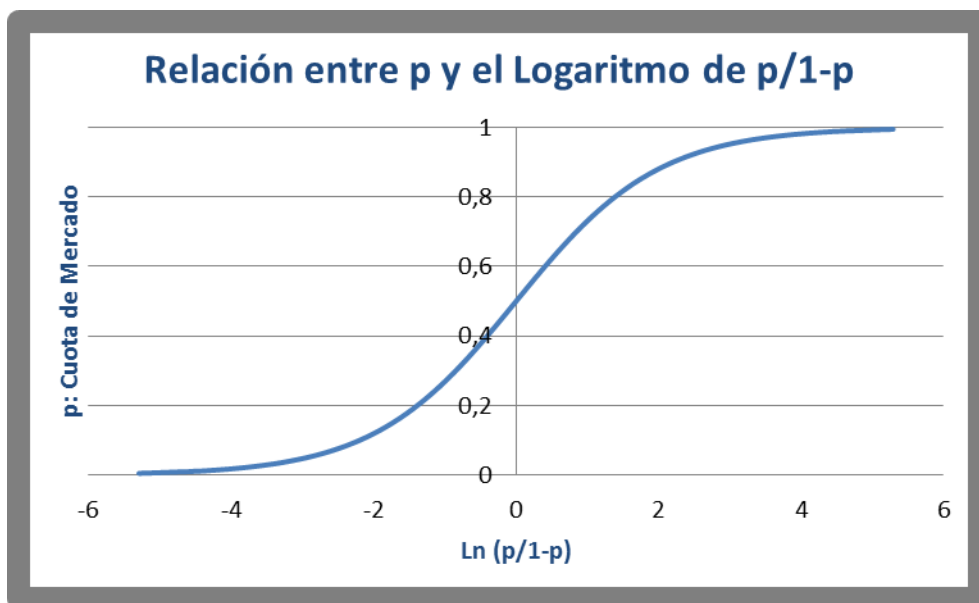
Se han elegido estos modelos porque estudian los efectos de variables de marketing mix sobre las ventas y porque aplican métodos econométricos al área de Marketing.

Ahora bien, pero la pregunta más adecuada debería ser ¿qué tienen en común y en que se diferencian estos modelos con el que se presenta en esta tesis y por tanto qué aporta?

La primera analogía se produce respecto a los modelos de publicidad. En estos modelos, la función de respuesta utilizada para especificar la relación entre las ventas y la publicidad, es cóncava o tiene forma de S. Esta forma funcional tiene sentido, ya que es esperable que en un primer momento, un ligero aumento de la publicidad produzca un aumento notable de las ventas, pero cuando el mercado se va saturando, el aumento de la publicidad se ve correspondido con ligeros aumentos de las ventas. El modelo que se presenta tiene como variable a explicar

$Ln\left(\frac{p}{1-p}\right)$  donde  $p$  es la cuota de mercado en tienda. La variable muestra la relación existente entre la cuota de mercado de la marca  $p$ , frente a la cuota de mercado de la competencia  $1-p$ , para posteriormente calcular el logaritmo neperiano de esta

relación. La ventaja de esta variable es que la forma en S buscada por los modelos publicitarios está garantizada y cumple con las restricciones impuestas por la realidad, ninguna cuota de mercado puede ser negativa o superior a 1.



Tiene una ventaja adicional y es que pueden introducirse todas las variables explicativas en una función lineal, por lo que su estimación será más sencilla que en otras formas funcionales no lineales.

La segunda coincidencia se produce respecto al tipo de modelo. El modelo planteado en la tesis es un modelo tipo SCAN\*PRO. Y es un modelo SCAN\*PRO porque utiliza datos de un panel de establecimientos proporcionados en este caso por Nielsen. Ejemplos de este tipo de modelos los encontramos en Wittink et al's (1988)

$$Ventas_{it} = \left( \prod_{j=1}^B (\text{Precio}_{jt})^{\beta_{ij}^{\text{Precio}}} \right) \left( \prod_{j=1}^B (\text{Publicidad}_{jt})^{\beta_{ij}^{\text{Publicidad}}} \right) \left( \prod_{k=1}^K (X_{ikt})^{\beta_{ik}} \right) \exp(u_{it})$$

Y en el modelo de Van Heerde y Neslin (Wierenga 2008:124).

$$S_{bst} = \lambda_{bs} \mu_{bt} \prod_{b'=1}^B \left\{ PI_{b'st}^{\beta_{b'b}} \cdot \gamma_{1b'b}^{FEATONLY_{b'st}} \cdot \gamma_{2b'b}^{DISPONLY_{b'st}} \cdot \gamma_{3b'b}^{FEAT \& DISP_{b'st}} \right\} e^{u_{bst}}$$

En ambos casos se tratan de modelos SCAN\*PRO, el primero mide la influencia del precio y la publicidad en las ventas de la marca en los establecimientos. El segundo la influencia del precio y de otras variables del marketing (la existencia de promoción, la existencia de display y la interacción entre promoción y display) sobre las ventas de la marca en los puntos de venta. Estos modelos y el presentado en la tesis comparten el tipo de dato; datos que muestran la evolución de las ventas de las marcas en los distintos establecimientos, durante un periodo determinado.

Ahora bien, las diferencias son notables si consideramos la forma funcional y las variables incluidas en los modelos. El modelo presentado trabaja con un número mayor de variables, recoge la influencia de las acciones de la competencia y analiza con mayor detalle el tipo de promoción realizado por la marca y la intensidad de la promoción (importancia de los establecimientos que han participado en la promoción). Respecto a la variable “Display”, entendida como la forma de anunciar la promoción, el modelo la recoge dividiéndola en dos variables diferentes: Folleto y Cabecera de góndola, permitiendo su análisis por separado y de forma conjunta (interacción), como aparece en el modelo de Van Heerde y Neslin (Wierenga 2008:124). Sólo añadir que el tratamiento de la variable precio se realiza con mayor detalle, no se construye un índice único como es el caso de estos modelos, sino diferentes índices para cada marca competidora.

La tercera analogía se produce con los modelos de Gestión de Ventas. Estos modelos utilizan una forma funcional que garantiza que la variable a explicar tenga forma de S y además introducen como variables explicativas otras variables distintas a la publicidad y la promoción. El modelo objeto de esta tesis presenta la misma forma funcional, con la cuota de mercado, pero simplificando su estimación, ya que veremos que se estima linealmente e introduce variables relacionadas con la distribución del producto, por lo tanto, variables no exclusivamente publicitarias o promocionales como hacen estos modelos de Gestión de Ventas.

Si consideramos a los modelos de Respuesta Competitiva, la analogía se produce porque en ambos casos se introducen como variables explicativas, variables que recogen las acciones de marketing realizadas por la competencia. Se entiende que la cuota de mercado de una marca es fruto de las acciones de la propia marca y de

las acciones de sus competidores, si esto es así, es lógico incluir esas acciones de la competencia en el modelo. Como se comprobará posteriormente, estas acciones de la competencia resultan ser estadística y económicamente significativas, reforzando la idea de su inclusión.

Por último, respecto a los modelos de series temporales, las semejanzas se producen porque ambos incluyen variables promocionales tanto propias como de la competencia. Otra similitud importante está en la utilización de variables retardadas, se trata en ambos casos de modelos dinámicos. Las grandes diferencias están en el tipo de dato que utilizan; mientras que los modelos de serie temporal, trabajan con datos desagregados a nivel de establecimiento, el modelo de esta tesis trabaja con datos agregados (globalmente, por formato o por tipo de establecimiento) y en el método de estimación, los modelos de series temporales permiten estimar un modelo para cada establecimiento, pero tal y como Leeflang (2008) reconoce, no hay un modelo único, sino un modelo para cada establecimiento. Por tanto cuando se trata de aplicar y mantener el modelo en la práctica, resulta muy complejo. El modelo planteado utiliza el método de estimación con datos de panel; que como se verá, permite rechazar la hipótesis de un modelo único, como el caso anterior, pero si permite una agrupación, ya sea por tipo de establecimiento, ya sea por formato. En la práctica esto se traduce en modelos más sencillos de mantener y de aplicar a situaciones concretas empresariales.

En el capítulo 2 se aborda la metodología que se utiliza en la tesis. Esta metodología no es otra que la regresión con datos de panel. La metodología de análisis viene determinada, en este como en otros, por el tipo de dato disponible. Los datos utilizados en esta tesis han sido proporcionados por Nielsen; y son los datos habituales que esta compañía proporciona a sus clientes. Son datos agregados que muestran la evolución en el tiempo de diferentes magnitudes de marketing. Esta agregación puede realizarse por formato, por tipo de establecimiento e incluso por zona de ventas.

El nivel de agregación de los datos hace imposible utilizar metodologías como las expuestas en modelos de series temporales mencionados anteriormente, ya que

sólo es posible aplicarla cuando se tiene información desagregada para cada uno de los establecimientos; y desgraciadamente este no es el caso.

Por otra parte, la aplicación de la metodología de series temporales supone el problema determinar si existe un modelo para cada establecimiento o bien un modelo único y global. En el primer caso, un modelo único, supone admitir que la realidad comercial es idéntica independientemente del formato que se comercialice o el establecimiento donde se venda; una realidad demasiado simple. El otro caso supone admitir un modelo diferente para cada formato y establecimiento; una realidad compleja, ya que desde la óptica de la gestión empresarial implica una mayor dificultad en el momento de analizar la realidad comercial.

En consecuencia, la pregunta es obvia, ¿existe algún punto intermedio entre ambas hipótesis?, ¿existe alguna técnica que nos permita encontrar puntos en común entre formatos y establecimientos y haga posible formular modelos más adaptados a las diferentes realidades comerciales pero sin estar subyugados por la necesidad de un modelo para cada establecimiento?

Es aquí donde los modelos de datos de panel se han mostrado sumamente útiles, mostrando evidencias respecto a la existencia de un único modelo, hipótesis que ha sido imposible de mantener. O bien, proporcionando una guía para agrupar los modelos, de manera que sin llegar a tener un modelo para cada establecimiento o producto si ha podido encontrarse un punto en común entre el análisis de la heterogeneidad comercial y el compromiso con la aplicación y mantenimiento de los modelos.

El capítulo continúa con la revisión de los métodos de estimación de datos de panel y que muchos de ellos se han utilizado en esta tesis. En un primer paso se define el Modelo Lineal de Datos de Panel  $y_{it} = x_{it}\beta + u_{it}$   $i = 1, \dots, N$ ,  $t = 1, \dots, T$  donde el subíndice  $i$  denota al individuo y el subíndice  $t$  denota el momento tiempo al que está referida la observación del individuo.

La ventaja o el inconveniente que presentan son los diferentes métodos de estimación del parámetro; así en capítulo estudiamos:

- El Método MCO Agrupado; método que no tiene en cuenta la información que proporcionan las diferencias entre individuos y utiliza exclusivamente su evolución en el tiempo.  $y_t = x_t\beta + u_t \quad t=1,...,T$ . Las variables explicativas son tratadas como un todo, pero despreciando la información de a qué individuo pertenece cada una de las variables. Es el método más parecido a MCO. El

estimador viene dado por la expresión:  $\hat{\beta} = \left( \sum_{i=1}^N \sum_{t=1}^T x'_{it} x_{it} \right)^{-1} \left( \sum_{i=1}^N \sum_{t=1}^T x'_{it} y_{it} \right)$

- Modelos de Datos de Panel con efectos no observados. En ocasiones, el motivo de usar datos de panel es resolver el problema de variables omitidas. Variables omitidas, constantes en el tiempo, y que pueden explicar las diferencias entre individuos, empresas, etc. Estas variables omitidas o efectos no observados serán tratados como VARIABLES ALEATORIAS. Ahora bien, el tema realmente importante es si los efectos no observados están incorrelacionados o no con las variables explicativas. El modelo básico de efectos no observados puede escribirse como  $y_{it} = x_{it}\beta + c_i + u_{it} \quad t=1,2,...,T \quad i=1,2,...,N$  en el que  $c_i$  es la variable que recoge los efectos no observados. Este efecto no observado ha generado toda una discusión metodológica respecto si debe ser tratado como un efecto aleatorio o como un efecto fijo. En la econometría moderna, la discusión está superada pero los nombres se mantuvieron, así cuando se habla de efectos aleatorios quiere decir que no hay correlación las variables explicativas y los efectos no observados. Cuando se habla de efectos fijos quiere decir que se está permitiendo correlación arbitraria entre los efectos no observados y las variables explicativas. Por tanto aparecen dos estimadores diferentes: El estimador de

Efectos Aleatorios  $\hat{\beta}_{EA} = \left( \sum_{i=1}^N X'_i \hat{\Omega}^{-1} X_i \right)^{-1} \left( \sum_{i=1}^N X'_i \hat{\Omega}^{-1} Y_i \right)$  que pasa por obtener  $\hat{\Omega}$ , y

el estimador de Efectos Fijo  $\hat{\beta}_{FE} = \left( \sum_{i=1}^N \ddot{X}'_i \ddot{X}_i \right)^{-1} \left( \sum_{i=1}^N \ddot{X}'_i \ddot{Y}_i \right) = \left( \sum_{i=1}^N \sum_{t=1}^T \ddot{x}'_{it} \ddot{x}_{it} \right)^{-1} \left( \sum_{i=1}^N \sum_{t=1}^T \ddot{x}'_{it} \ddot{y}_{it} \right)$

conocido también con el nombre de estimador WITHIN.

- Método de Primeras Diferencias. Partiendo del modelo utilizado hasta el momento  $y_{it} = x_{it}\beta + c_i + u_{it}$   $t=1,2,\dots,T$ ;  $i=1,2,\dots,N$  la transformación en diferencias consiste en tomar el modelo original y retardarlo un periodo; a continuación se resta uno de otro y llegamos a la expresión:  $\Delta y_{it} = \Delta x_{it}\beta + \Delta u_{it}$   $t=2,3,\dots,T$  donde  $\Delta y_{it} = y_{it} - y_{i,t-1}$ ,  $\Delta x_{it} = x_{it} - x_{i,t-1}$  y por último  $\Delta u_{it} = u_{it} - u_{i,t-1}$ . Esta transformación elimina el efecto no observado  $c_i$ . Al tomar diferencias se pierde el primer periodo de cada sección cruzada  $i$ , de manera que nos quedamos con  $T-1$  periodos temporales para cada  $i$  en lugar de los  $T$  que teníamos antes. La ecuación  $\Delta y_{it} = \Delta x_{it}\beta + \Delta u_{it}$   $t=2,3,\dots,T$  pone de manifiesto que los elementos a incluir en  $x_{it}$  deben variar en el tiempo, en caso contrario serán eliminados en la transformación. El estimador de primeras diferencias (PD),  $\hat{\beta}_{PD}$  es el estimador MCO Agrupados de la regresión  $\Delta y_{it}$  sobre  $\Delta x_{it}$   $t=2,3,\dots,T$ ,  $i=1,2,\dots,N$
- En la última parte del capítulo se aborda la comparación entre los estimadores de efectos fijos y primeras diferencias y la relación entre los estimadores de efectos aleatorios y efectos fijos. Demostrando que el estimador de Efectos Aleatorios es comparable con el de Efectos Fijos. El elemento clave aparece en la transformación  $\tilde{y}_i$  que es  $y_{it} - \lambda \bar{y}_i$  y lo mismo para  $\tilde{X}_i$ . El estimador de Efectos Aleatorios factible reemplaza el desconocido  $\lambda$  por su estimador  $\hat{\lambda}$ , por lo que  $\hat{\beta}_{EA}$  puede calcularse de la regresión de MCO Agrupados  $\tilde{y}_{it}$  sobre  $\tilde{x}_{it}$   $t=1,2,\dots,T$ ,  $i=1,2,\dots,N$  donde  $\tilde{x}_{it} = x_{it} - \hat{\lambda} \bar{x}_i$  y  $\tilde{y}_{it} = y_{it} - \hat{\lambda} \bar{y}_i$ . Por tanto el estimador de efectos aleatorios puede escribirse como

$$\hat{\beta}_{EA} = \left( \sum_{i=1}^N \sum_{t=1}^T \tilde{x}_{it}' \tilde{x}_{it} \right)^{-1} \left( \sum_{i=1}^N \sum_{t=1}^T \tilde{x}_{it}' \tilde{y}_{it} \right).$$

Esta expresión se obtiene por una “QUASITIME DEMEANING”. En lugar de eliminar totalmente el promedio temporal en las variables explicativas y en la variable dependiente, los efectos aleatorios eliminan una fracción de dicho promedio temporal. Si  $\hat{\lambda}$  está cerca de la unidad, las estimaciones de efectos aleatorios y las de efectos fijos tienden a estar cerca. Para verlo escribimos  $\hat{\lambda}$  como:

$$\hat{\lambda} = 1 - \left[ \frac{1}{1 + T \left( \frac{\hat{\sigma}_c^2}{\hat{\sigma}_u^2} \right)} \right]^{\frac{1}{2}} \quad \hat{\sigma}_u^2 = \hat{\sigma}_v^2 - \hat{\sigma}_c^2, \quad \hat{\sigma}_v^2 = \frac{1}{(NT - K)} \sum_{i=1}^N \sum_{t=1}^T \hat{v}_{it}^2, \text{ siendo } \hat{v}_{it} \text{ los residuos}$$

de la regresión de MCO Agrupados y  $\hat{\sigma}_c^2 = \frac{1}{\left[ N \frac{T(T-1)}{2} - K \right]} \sum_{i=1}^N \sum_{t=1}^{T-1} \sum_{s=t+1}^T \hat{v}_{it} \hat{v}_{is}$ . Cuando

$T \left( \frac{\hat{\sigma}_c^2}{\hat{\sigma}_u^2} \right)$  es grande, el segundo término de  $\hat{\lambda}$  es pequeño y en estos casos  $\hat{\lambda}$

está cerca de la unidad. En efecto,  $\hat{\lambda} \rightarrow 1$  cuando  $T \rightarrow \infty$  o cuando  $\left( \frac{\hat{\sigma}_c^2}{\hat{\sigma}_u^2} \right) \rightarrow \infty$ .

Para una  $T$  grande no es sorprendente encontrar estimaciones similares de efectos fijos y de efectos aleatorios. Incluso con  $T$  pequeña los efectos aleatorios pueden estar cerca de los efectos fijos si la estimación de la varianza de  $c_i$  es relativamente grande respecto a la varianza estimada de  $u_{it}$ . Cuando  $\lambda$  se aproxima a la unidad, la precisión del estimador de efectos aleatorios se aproxima al estimador de efectos fijos, y los efectos de las variables explicativas constantes en el tiempo se vuelven difíciles de estimar. Las expresiones

$$\hat{\beta}_{EA} = \left( \sum_{i=1}^N \sum_{t=1}^T \tilde{x}'_{it} \tilde{x}_{it} \right)^{-1} \left( \sum_{i=1}^N \sum_{t=1}^T \tilde{x}'_{it} \tilde{y}_{it} \right) \quad \text{y} \quad \hat{\lambda} = 1 - \left[ \frac{1}{1 + T \left( \frac{\hat{\sigma}_c^2}{\hat{\sigma}_u^2} \right)} \right]^{\frac{1}{2}} \text{ muestran también la}$$

relación existente entre los efectos aleatorios y los MCO Agrupados. Los MCO Agrupados se obtienen haciendo  $\hat{\lambda} = 0$ , que nunca es exactamente cierto pero podría estar cerca. En la práctica  $\hat{\lambda}$  no está habitualmente cerca de 0 porque este resultado requeriría que  $\hat{\sigma}_u^2$  fuera relativamente grande respecto a  $\hat{\sigma}_c^2$ .

Por último se realiza una primera aproximación al Test de Hausman, contraste necesario para comparar los estimadores de Efectos Fijos y Efectos Aleatorios. Esta aproximación se realiza desde la óptica de su utilidad, ya que en el capítulo siguiente tendrá un tratamiento más detallado.



El elemento clave para elegir entre ambos estimadores es saber si  $c_i$  y  $x_{it}$  están correlacionados, Hausman (1978) propuso un test basado en las diferencias entre las estimaciones de efectos aleatorios y efectos fijos. El estimador de Efectos Fijos es consistente cuando  $c_i$  y  $x_{it}$  están correlacionados, mientras que el de Efectos Aleatorios es inconsistente; por lo tanto una diferencia estadísticamente significativa se interpreta como una evidencia en contra de la hipótesis de efectos aleatorios. Es decir, la hipótesis de que  $c_i$  y  $x_{it}$  están incorrelacionados no puede mantenerse y lo adecuado en ese caso sería utilizar el estimador de Efectos Fijos.

El capítulo 3 se centra en revisar los contrastes que más adelante se irán necesitando. El capítulo no es una recopilación exhaustiva de todos los contrastes existentes, más bien, es una recopilación interesada. Se recogen los contrastes que tienen más posibilidades de utilizarse en los capítulos posteriores.

El capítulo comienza con el problema de la heteroscedasticidad y la autocorrelación (detección y corrección), recogiendo un planteamiento alternativo, como es la utilización de estimadores robustos. Este es el cuerpo principal del capítulo, dada la gran cantidad de material disponible que trata estos problemas. Posteriormente y de una forma más detallada, volvemos sobre el contraste de Hausman. A continuación repasamos los test de Agrupabilidad, estos contrastes se utilizan para determinar si la especificación de un modelo de datos de panel tiene sentido, cuestión absolutamente capital cuando se pretende estimar un modelo de datos de panel. Y para finalizar el contraste de especificación RESET. Este contraste permite determinar la existencia de no linealidades no recogidas en la estructura del modelo y por tanto incluirlas en él. El desarrollo teórico de este contraste se encuentra fundamentado en modelos de corte transversal. Aquí puede verse su aplicación en datos de panel, aplicación inédita en este campo, de la que me siento particularmente satisfecho.

El capítulo comienza con la formulación de dos problemas clásicos en la Econometría: La heteroscedasticidad y la autocorrelación, problemas que también aparecen en los datos de panel. El esquema utilizado es un esquema clásico; se analiza cada problema por separado, la forma de detectarlos y corregirlos.

- **HETEROSCEDASTICIDAD.** En el modelo de datos de panel:  
 $y_{it} = x_{it}\beta + v_{it} \quad t = 1, 2, \dots, T$  donde  $v_{it} = c_i + u_{it} \quad t = 1, 2, \dots, T$  se asume que perturbación aleatoria es homoscedástica; todos los individuos tienen la misma varianza y se mantiene constante en el tiempo. Esta hipótesis es bastante restrictiva para datos de panel, ya los individuos pueden cambiar de tamaño y como resultado tener varianzas. Asumir que la perturbación aleatoria es homoscedástica cuando no lo es, produce estimaciones consistentes de los parámetros del modelo, pero estas estimaciones no serán eficientes; su varianza no será mínima. Además los errores estándar de dichas estimaciones serán sesgados.

Los contrastes a estudiar varían esencialmente en el tratamiento dado al individuo (unidad de sección cruzada). Se busca la presencia de heteroscedasticidad en cada individuo y posteriormente se comprueba si la varianza en cada uno de ellos es la misma. Estos contrastes no son muy habituales en Econometría, ni tampoco en Econometría de Datos Panel, entonces ¿por qué estudiarlos? La razón está en la tipología de datos utilizados en este trabajo. Los datos utilizados se caracterizan porque el número de individuos “N” es muy pequeño, mientras que el tiempo “T” es elevado, de tal modo que se cumple que  $T > N$ . Esta situación no es habitual en datos de panel, donde lo normal es tener un número elevado de individuos, pero durante un periodo pequeño de tiempo. La principal aportación de estos contrastes es que están diseñados para comparar la varianza entre individuos y concluir que todos tienen la misma varianza o no. Los contraste de homoscedasticidad utilizados habitualmente en econometría, buscan el cumplimiento de la homoscedasticidad dentro del individuo, por lo tanto no tienen sentido contrastes como los anteriores. O si existen individuos, como en los datos de panel, el número de observaciones disponible de cada individuo es muy limitado por que lo que la aplicación práctica de estos contraste es complicada. Los contrastes para detectar la presencia de heteroscedasticidad son:

- Contraste de Heteroscedasticidad basado en el test de Bartlett. Este contraste comprueba la existencia de heteroscedasticidad dentro de cada individuo. Los datos de cada individuo son datos temporales, referidos al mismo individuo y tal y como indican los manuales de econometría, no suelen presentar problemas de heteroscedasticidad. Para comprobar que efectivamente es así, puede utilizarse el test de White, en la versión propuesta por Wooldridge, (2005: 298). Una vez aceptada la hipótesis nula de homoscedasticidad en cada individuo, se comprueba si las varianzas de los distintos individuos son iguales; para ello puede utilizarse el test de Bartlett.
- Contraste de Heteroscedasticidad basado en el test de Levene. Así como el test de Bartlett aparece recogido en algunos manuales de econometría, no sucede lo mismo con el test de Levene. Este contraste tiene la característica de ser menos sensible respecto al incumplimiento de la hipótesis de normalidad.
- Corrección de la Heteroscedasticidad. Detectada la presencia de heteroscedasticidad es preciso transformar el modelo. Se estudia la transformación propuesta por Baltagi, (2001:78). Esta transformación supone una modificación personalizada del modelo para cada individuo, de manera que dos individuos del mismo panel sufrirán distintas transformaciones. Cuando se pone en práctica la transformación (MCG Factibles) queda claro inmediatamente que este procedimiento requiere una  $T$  grande y una  $N$  preferiblemente pequeña con  $T > N$ . Esta no es la típica situación de datos de panel pero si el caso de nuestro trabajo donde nos encontramos sólo con 4 individuos y más de 100 observaciones en el tiempo de cada uno de ellos.

- **AUTOCORRELACIÓN.** La estimación en presencia de autocorrelación en el término de la perturbación aleatoria produce estimaciones consistentes pero ineficientes y los errores estándar de dichas estimaciones, de nuevo estarán sesgados. El problema es de tal índole que no queda más remedio que corregir el modelo; ya que en caso contrario el trabajo realizado no sería válido.

Para detectar la presencia de autocorrelación nos enfrentamos al problema de contrastar conjuntamente autocorrelación y efectos individuales. Baltagi y Li (1995) derivaron tres estadísticos LM para un modelo de componentes del error, con errores con autocorrelación de orden 1. En todos estos casos Baltagi y Li (1995) demostraron que el estadístico LM es el mismo tanto si la hipótesis alternativa es un AR(1) o un MA(1). Estos contrastes son simples en el sentido de que sólo requieren la obtención de residuos MCO. En este capítulo se realiza un estudio detallado de estos contrastes, concretamente:

- Test LM conjunto para Autocorrelación y Efectos Individuales Aleatorios. Baltagi (2001:90) propuso un test LM conjunto para un modelo de componentes del error donde la distribución sigue un proceso MA(1). La hipótesis a contrastar es  $H_1^a : \sigma_\mu^2 = 0; \lambda = 0$ . Baltagi y Li (1991b) demostraron que este contraste es exactamente el mismo que el contraste necesario para testar la hipótesis:  $H_1^b : \sigma_\mu^2 = 0; \rho = 0$  en el que la perturbación aleatoria sigue un proceso AR(1). Por lo que el test estadístico conjunto LM para efectos individuales aleatorios y autocorrelación de primer orden es independiente de la forma de la autocorrelación, tanto si es un AR(1) o un MA(1).
- Test LM conjunto para Autocorrelación de primer orden en un modelo de Efectos Aleatorios. Baltagi y Li (1995) derivaron un test LM para autocorrelación de primer orden, suponiendo que existan efectos individuales aleatorios. La hipótesis nula es:

$$H_4^b : \rho = 0 \quad \left( \text{dado } \sigma_\mu^2 > 0 \right)$$

$$H_4^{b'} : \rho \neq 0 \quad \left( \text{dado } \sigma_\mu^2 > 0 \right)$$

Baltagi y Li llegaron a la conclusión de que este test era exactamente el mismo que en el caso anterior. Por tanto y para resumir, el test LM para contrarautocorrelación de primer orden, asumiendo efectos individuales aleatorios, es invariable a la forma de la autocorrelación (si es AR(1) o MA(1)). Más adelante, las aportaciones del propio Baltagi (2001: 93), Bera, Sosa – Escudero y Yoon (2001), lo convierten en robusto frente a un problema de especificación local.

- Test LM conjunto para Autocorrelación de primer orden en un modelo de Efectos Fijos. Dado el modelo  $y_{it} = z'_{it}\delta + u_{it} \quad i=1\dots N \quad t=1\dots T$  en el que  $\delta$  es un vector de  $(k+1) \times 1$  coeficientes incluyendo la constante  $u_{it} = \mu_i + v_{it}$ . La hipótesis nula es  $H_5^b : \rho = 0$ .
- Test Durbin – Watson para datos de panel. Para el modelo de efectos fijos  $y_i = z_i\delta + \mu_i + v_i$  con  $v_i$  siguiendo un proceso AR(1), Bhargava, Franzini y Narendranathan (1982) sugirieron un test  $H_o : \rho = 0$  frente a la alternativa  $|\rho| < 1$ , utilizando el estadístico de Durbin – Watson basado en los residuos WITHIN ( $\tilde{v}_{it}$ ) en lugar de los residuos MCO. Los autores demostraron que para regresores arbitrarios, el estadístico utilizado era un test más poderoso para valores cercanos a  $\rho = 0$ . Los mismos autores también nos sugieren utilizar el estadístico Berenblut – Webb para testar  $H_o : \rho = 0$  porque es un test invariante localmente más poderoso para valores cercanos a  $\rho = 1$ .  
Bhargava, Franzini y Narendranathan también proponen un test para comprobar si los residuos siguen un Paseo Aleatorio; esto es,  $H_o : \rho = 1$  frente a la alternativa  $|\rho| < 1$ .
- Contraste para diferenciar entre un AR(1) frente a un MA(1) en un modelo de componentes del error. Burkey, Godfrey y Termayne (1990) derivaron este contraste para distinguir entre ambos procesos. La idea básica es que bajo la hipótesis nula de un proceso AR(1), el error  $v_{it}$

satisface la condición:  $correl(v_{i,t}, v_{i,t-\tau}) = \rho_\tau = (\rho_1)^\tau \quad \tau = 1, 2, \dots$  por lo tanto bajo la hipótesis nula, se tiene que cumplir que  $\rho_2 - (\rho_1)^2 = 0$ . Bajo la hipótesis alternativa de un proceso MA(1) sobre  $v_{it}$ ,  $\rho_2 = 0$  y por tanto  $\rho_2 - (\rho_1)^2 \leq 0$ . Posteriormente, Baltagi y Li (1995) propusieron una modificación de este contraste BGT.

- Contraste alternativo tipo BGT para contrastar un AR(1) frente un MA(1). Este test, tipo BGT, utiliza estimaciones de autocorrelación derivadas de residuos MCO. Baltagi y Li (1995) demostraron que para  $N$  grande es posible distinguir entre el proceso AR(1) y el MA(1). Sin embargo, los mismos autores, realizaron experimentos Monte Carlo, y la distribución resultante no permitía distinguir si se trataba de un proceso AR(1) o MA(1).

En el capítulo se presentan las conclusiones más importantes respecto a estos contrastes.

- **CORRECCIÓN DE LA AUTOCORRELACIÓN.** En primer lugar se estudia el caso de autocorrelación generada por un AR(1) en un modelo de Efectos Aleatorios. En el capítulo se trabaja de manera detallada con la transformación propuesta por Baltagi y Li (1991a) que está derivada de la transformación de Fuller y Battese (1974). Además se compara esta transformación con la transformación  $\lambda$  propuesta por Wooldridge (2002; 287). Se llega a la conclusión de que se trata de la misma transformación.

En páginas posteriores se expone la corrección de la autocorrelación generada por un proceso AR(1) en el modelo de Efectos Fijos. Este procedimiento fue desarrollado por Bhargava, Franzini y Narendranathan (1982).

- **UN PLANTEAMIENTO ALTERNATIVO.** Cada vez es más habitual, especialmente en los trabajos de carácter práctico, utilizar estimadores que

sean robustos a, bien a heteroscedasticidad, bien a autocorrelación e incluso robustos a la presencia de ambos problemas. La ventaja resulta evidente, no tenemos que preocuparnos de como corregir los problemas, y nos centramos exclusivamente en la significatividad de las variables introducidas en el modelo, aunque si bien es cierto, siempre a costa de perder precisión. En datos de panel, también están disponibles estimadores de este tipo. Más concretamente, en este trabajo se han utilizado “cluster robust standard error” (errores robustos por cluster o conglomerado). Carter Hill (2012:85) nos muestra, a través de un modelo de datos de panel muy sencillo, el funcionamiento de estos errores robustos por cluster.

- **TEST DE ESPECIFICACIÓN DE HAUSMAN.** En el capítulo se vuelve sobre este contraste que se había introducido en el capítulo anterior, para tratarlo con mayor detalle. La cuestión primordial es ¿con qué modelo debemos trabajar, con el de Efectos Fijos o con el de Efectos Aleatorios? Para responder a la pregunta es preciso comprobar si se cumple una hipótesis básica en el modelo de componentes del error. Esto es, si se cumple que  $E[u_{it}/X_{it}] = 0$ . Si  $E[u_{it}/X_{it}] \neq 0$  el estimador  $\hat{\beta}_{MCG}$  que usamos en el modelo de Efectos Aleatorios se convierte en un estimador SESGADO e INCONSISTENTE de  $\beta$ . Por otra parte, con la transformación Within eliminamos los  $\mu_i$  y el estimador  $\hat{\beta}_{Within}$  es un estimador SESGADO y CONSISTENTE de  $\beta$ . Esta cuestión fue resuelta por Hausman (1978) comparando ambos estimadores:  $\hat{\beta}_{MCG}$  y  $\hat{\beta}_{Within}$ .

Bajo  $H_0: E[u_{it}/X_{it}] = 0$  ambos estimadores son consistentes; en el caso de que  $H_0$  sea falsa tendrán límites en probabilidad diferentes. Esto significa que el estimador  $\hat{\beta}_{Within}$  es siempre consistente, tanto si  $H_0$  es verdadera o falsa, mientras que el estimador  $\hat{\beta}_{MCG}$  es ELIO, consistente y asintóticamente eficiente bajo  $H_0$ , pero es INCONSISTENTE cuando  $H_0$  es falsa. El contraste estadístico natural debería basarse en comparar ambos estimadores:

$\tilde{q}_1 = \hat{\beta}_{MCG} - \hat{\beta}_{Within}$  Bajo  $H_0$ ,  $p \lim \hat{q}_1 = 0$  y  $cov(\hat{q}_1, \hat{\beta}_{MCG}) = 0$ . En el capítulo se realiza esta demostración.

- **TEST DE AGRUPABILIDAD.** El campo de aplicación de estos contraste es conocer si la modelización mediante de datos de panel tiene sentido. Hsiao (2002: 13) plantea tres modelos distintos a aplicar con la intención de elegir el modelo más adecuado a nuestros datos.

$$y_{it} = \alpha_i + \beta_i' x_{it} + u_{it}$$

Cada individuo tiene  $\alpha$  y  $\beta$  distinta. Esto significa que cada individuo tiene un modelo específico; es el caso menos restrictivo posible.

$$y_{it} = \alpha_i + \beta' x_{it} + u_{it}$$

Este modelo contempla un  $\alpha$  distinto para cada individuo. Se trata del modelo de Efectos Individuales Específicos que puede ser estimado por los métodos de Efectos Fijos, Efectos Aleatorios, Primera Diferencia, etc.

$$y_{it} = \alpha + \beta' x_{it} + u_{it}$$

Este modelo significa que todos los parámetros son iguales para cada individuo. Es el modelo MCO AGRUPADO y es el modelo más restrictivo posible.

Los contrastes enunciados son contraste de comparación de suma de residuos al cuadrado, propuestos por Chow (1960) y utilizados muy habitualmente en otras áreas de la econometría. Los estadísticos de contraste resultantes son estadísticos F ya que plantean un cociente de sumas de cuadrados y cada una de ellas sigue una chi cuadrado.

El objetivo perseguido es determinar qué modelo de datos de panel debemos utilizar, así, si concluimos que los parámetros son iguales para todos los individuos, el modelo idóneo es el Datos de Panel Agrupados. Si esta hipótesis resulta rechaza, se deberá decir si el modelo idóneo es el que tiene



un intercepto común y pendiente distinta para los distintos individuos o el modelo más general que presenta intercepto y pendiente distinta para cada uno de los individuos.

- **CONTRASTE RESET (Regression Specification Error Test) O DE ERROR DE ESPECIFICACIÓN.** Carter Hill, Griffiths y Lim (2012:151) hacen referencia en su obra, al contraste RESET de Ramsey (1969). Este contraste fue propuesto por Ramsey para detectar variables omitidas y formas funcionales incorrectas. El autor, en el artículo original lo compara además con otros contrastes alternativos. Rechazar la hipótesis nula implica que el modelo original no es adecuado y puede ser mejorado. La filosofía general del test es: si podemos mejorar significativamente el modelo aumentando artificialmente su poder de predicción, entonces el modelo original no es adecuado. Este contraste ha sido adaptado al contexto de datos de panel, ya que no se ha encontrado su aplicación en otros textos. En el capítulo se presenta su funcionamiento.

En el capítulo 4 se aborda la especificación del modelo. Además se contrastan las dos primeras hipótesis planteadas en esta tesis. La primera ¿Es posible construir un modelo que explique el comportamiento comercial de las marcas en el punto de venta? Esta pregunta se contesta en este capítulo.

La segunda hipótesis, está relacionada con la primera; una vez construido el modelo, ¿es único? o por el contrario será necesario construir tantos modelos como formatos y establecimientos tengamos. La contestación a esta cuestión se realiza al final del capítulo, una vez determinado el modelo con el que se trabaja.

La variable a explicar es la cuota en tienda, esto es, la cuota de mercado en los establecimientos donde está presente la marca. Los datos utilizados han sido facilitados por Nielsen; son datos del mercado español de cacao soluble. Nielsen ha borrado cualquier referencia a las marcas que operan en el mercado para garantizar el anonimato de sus clientes.

¿Por qué la cuota en tienda y no la cuota de mercado? Cuando se trabaja con datos procedentes de un panel de detallistas, la evolución de un mercado se mide por dos variables:

- La distribución, número de establecimientos que trabajan el producto y
- La cuota en tienda, cuota de mercado en los puntos de venta donde está presente la marca.

Esto significa que un mercado crece porque hay más establecimientos que venden el producto, o porque los establecimientos que venden el producto, venden más, o bien es una conjunción de los dos factores. Trabajar con la cuota de mercado directamente supone añadir una fuente de heterogeneidad más a los datos; la distribución, por esta razón es preferible trabajar con una variable más homogénea y que proporcione información sobre los que está sucediendo en los puntos de venta.

La forma de calcularla es:

$$\text{Cuota en tienda marca } i \text{ formato } j = \frac{\text{Ventas en tienda marca } i \text{ formato } j}{\text{Ventas Totales de los establecimientos donde está presente la marca } i \text{ formato } j}$$

Además la cuota de mercado no deja de ser una entelequia. Cuando realizamos una compra, esa compra se materializa en un producto concreto, con un formato concreto y en un establecimiento concreto. Es después cuando analizamos el mercado cuando se habla de cuota de mercado, como la suma de las ventas de una marca dividida por la suma de las ventas de todas las marcas. Pero realmente las compras se materializan en productos, formatos y establecimientos concretos. En este sentido trabajar con la cuota de mercado supone trabajar con una variable más heterogénea que la cuota en tienda.

Por tanto la variable a explicar: Cuota en tienda de la marca  $i$  será tratada como una variable de panel con la siguientes elementos:

- Los elementos del panel se representan con la letra  $i$  indicando los distintos formatos de la marca a analizar en los distintos tipos de establecimientos.

- Dentro de cada elemento encontramos 108 observaciones ordenadas en el tiempo con una periodicidad semanal. La primera observación corresponde a la semana 17.

Se introduce una variable ficticia. La estimación por datos de panel exige que las X se mantengan constantes a lo largo de toda la muestra; esto es, que estén presentes en todos los individuos y a lo largo de todo el periodo considerado. Esto supone que todas las marcas con todos sus formatos estén presentes en todos los establecimientos y a lo largo de todos los periodos. Imposible, en la información disponible encontramos casos donde una marca está presente en un tipo de establecimiento pero no en otro, y si está presente en un formato no significa que deba estarlo en todos los formatos considerados. Para resolver la situación se definen a las variables explicativas respecto al **PRINCIPAL COMPETIDOR**. Esto significa que en unos establecimientos o formatos el principal competidor puede ser la marca 4 y en otros puede serlo la marca 5 y sin embargo sus acciones aparecerán de forma conjunta en el análisis, como si se tratara de una sola marca.

Las consecuencias de esta decisión es que si bien posibilita la utilización de la metodología de datos de panel, producirá un modelo más simplista y más alejado de la realidad en los establecimientos más grandes donde el nivel de competencia es más elevado, mientras que en los establecimientos pequeños estará más próxima a la realidad.

Además se tienen en cuenta como variables explicativas en el análisis: **RESTO DE FABRICANTES Y MARCAS DEL DISTRIBUIDOR**, estas variables poseen una característica común con el **COMPETIDOR PRINCIPAL**, son fruto de la agregación de distintas marcas. Así el caso de Resto de Fabricantes está constituido por pequeños fabricantes cuya importancia individual es pequeña y son agrupados en una rúbrica más importante.

La MARCA DEL DISTRIBUIDOR recoge aquellos productos comercializados por las distintas cadenas con su marca privada. Como no existe una única cadena de establecimientos, nos encontramos con distintas marcas del distribuidor; por lo tanto, la MARCA DEL DISTRIBUIDOR constituye una agregación de distintas marcas de distribuidores.

El objetivo perseguido es explicar la Cuota en tienda de la marca 1 en todos sus formatos, en un entorno competitivo formado por el PRINCIPAL COMPETIDOR, RESTO DE FABRICANTES Y MARCAS DEL DISTRIBUIDOR. Sobre este entorno se definen el resto de variables a utilizar como explicativas y su interpretación.

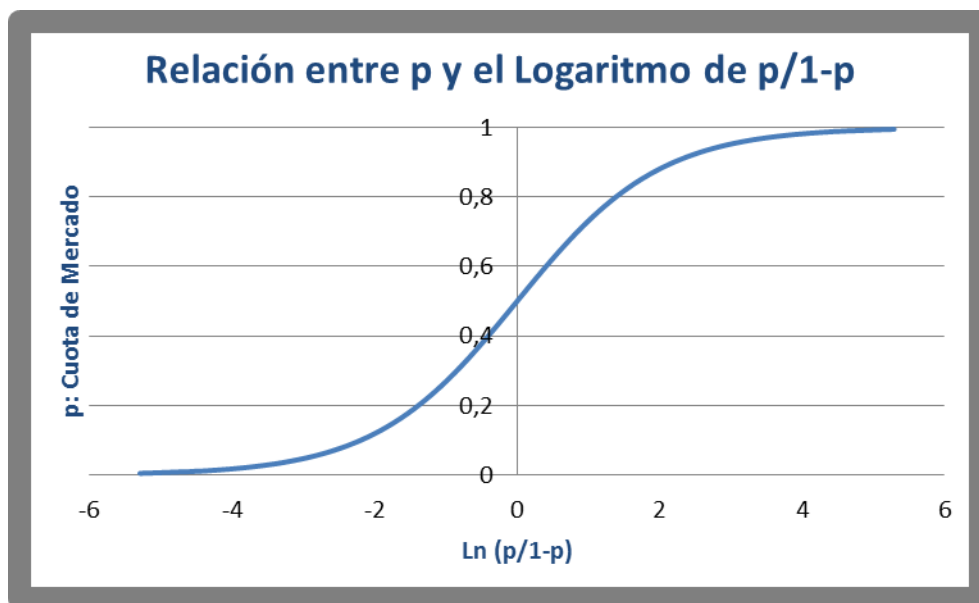
Estas variables se definen respecto al precio (se utiliza un índice de diferencial de precio que compara el precio de la marca con el de las restantes marcas), respecto a la Distribución Ponderada de las marcas (concepto desarrollado por Nielsen), respecto a la **intensidad** de la promoción de precio y respecto a la **cobertura** de los distintos tipos de promociones realizadas por la marca y recogidos por Nielsen. Estas dos últimas variables son una aportación original no utilizada hasta el momento en los modelos de marketing.

Una vez establecidas la variable a explicar y las posibles variables explicativas, los pasos siguientes van encaminados a la estimación del modelo. Pero en este caso nos encontramos con posibles modelos a estimar y diferentes métodos a utilizar.

En primer lugar se estima el modelo de forma estática, posteriormente se realiza un tratamiento dinámico de los datos.

Ahora bien, la estimación del modelo pone de manifiesto un problema que es necesario resolver previamente. El problema es que si sustituimos las variables explicativas por sus posibles valores llegaríamos a obtener valores en la variable a explicar que están fuera de su rango. La cuota en tienda podría tomar valores por debajo de 0 o por encima de 1; y esto, en la práctica, resulta imposible.

La solución al problema planteado la proporcionó el profesor Baltagi, en un curso que impartió en el mes de junio de 2010 en el Instituto de Estudios Fiscales en Madrid, y al que tuve la oportunidad de asistir. Baltagi propuso resolver el problema de la siguiente forma: en lugar de modelizar la cuota de mercado ( $p$ ) modelizar la relación  $\ln\left(\frac{p}{1-p}\right)$ ; a partir de esta relación se puede tratar linealmente su relación con las variables explicativas. Además, como ya se adelantaba en el primer capítulo, Leeflang y Wittink (2000) desarrollan una idea de Little respecto a la estructura de los modelos. Los modelos deben ser simples, completos respecto a las variables, adaptativos y robustos. La robustez puede lograrse con una estructura que restrinja las respuestas a un rango de valores que tenga sentido. Por tanto, si los valores de la variable explicativa están restringidos entre 0 y 1, como es el caso de la cuota de mercado, el modelo debería satisfacer estas mismas restricciones. Y podrían considerarse a este modelo como consistente. Es esta una razón más para incorporar la variable a explicar de acuerdo con la relación propuesta por Baltagi.



En el siguiente gráfico se muestra la relación existente entre  $p$  y  $\ln\left(\frac{p}{1-p}\right)$

Puede observarse que valores de  $p$  superiores al 50 % se corresponden con valores de  $\ln\left(\frac{p}{1-p}\right)$  positivos y valores de  $p$  inferiores al 50 % se corresponden con valores

negativos. Por otro lado, resulta imposible obtener en  $p$  valores superiores a 1 e inferiores a 0, con lo que el problema enunciado queda resuelto.

Resuelto el problema, se pasó a estimar el modelo por MCO Agrupados. El modelo a estimar es del tipo  $y_{it} = \alpha + \mathbf{x}'_{it}\beta + u_{it}$  que es modelo más restrictivo posible, establece que todos los coeficientes son iguales para los distintos individuos que componen el panel.

Si el modelo está correctamente especificado y los regresores incorrelacionados con el error, entonces puede estimarse CONSISTENTEMENTE por el método propuesto. El paso siguiente es comprobar si el modelo planteado es adecuado; esto es si el modelo tiene un intercepto común y una pendiente común para todos los individuos.

El contraste de agrupabilidad propuesto por Hsiao (2002: 15 - 21) nos proporciona la respuesta. El estadístico calculado para este modelo es el siguiente:

$$488,1539916 = F_4 \geq F_{(N-1), N(T-1)-K}^{(\alpha)} = F_{(24-1), 24(108-1)-36}^{(5\%)} = F_{23, 2532}^{(5\%)} = 0.971425$$

Y nos indica que debemos rechazar la hipótesis nula. Por lo tanto estamos en condiciones de afirmar que el modelo  $y_{it} = \alpha + \mathbf{x}'_{it}\beta + u_{it}$  no es el adecuado comparándolo con el modelo  $y_{it} = \alpha_i + \beta' x_{it} + u_{it}$ . Ya que el estimador utilizado resulta ser INCONSISTENTE.

Desde la óptica del marketing, la aplicación del contraste de Hsiao (2002), además de contrastar una de las hipótesis de partida contenidas en esta tesis, supone admitir que NO hay un único modelo de gestión para todos los formatos y establecimientos; sino que es necesario realizar adaptaciones a las particularidades de cada establecimiento y formato.

En este momento, se desconoce aún cuál será la profundidad de esta heterogeneidad. Esta heterogeneidad, presente en cada individuo, puede recogerse en una constante específica para él o por el contrario, es de naturaleza más amplia y será preciso recogerla en un modelo para cada individuo. O quizás quedarnos en un punto intermedio; esto es, un modelo para cada formato o un modelo para cada tipo de establecimiento.

Como ya sabemos que no existe un modelo único, es necesario estimar un modelo de Efectos Individuales Específicos. Este modelo  $y_{it} = \alpha_i + \mathbf{x}_{it}'\beta + u_{it}$  permite que cada individuo (unidad de sección cruzada) tenga un intercepto diferente, aunque la pendiente siga siendo la misma para todos los individuos. Los  $\alpha_i$  son variables aleatorias que recogen la heterogeneidad no observada.

En este modelo se puede tratar a  $\alpha_i$  como una variable aleatoria inobservada que potencialmente está correlacionada con los  $\mathbf{x}_{it}$  regresores observados. Este modelo recibe el nombre de modelo de efectos fijos (EF). Otra variante del modelo es asumir que los  $\alpha_i$ , que representan los efectos individuales inobservables, son variables aleatorias. En este caso hablamos del modelo de efectos aleatorios (EA).

La forma de proceder será estimar ambos modelos y después elegir el más adecuado para nuestro análisis. Las diferencias entre ambas estimaciones son bastante acusadas, pero para decidir con cuál de ellas debemos quedarnos utilizamos el Test de Hausman (1978). Los resultados obtenidos nos permiten rechazar la hipótesis nula; por tanto el estimador que debemos utilizar es el estimador de Efectos Fijos. En el capítulo se muestra una versión robusta de este contraste dada la gran diferencia existente entre las estimaciones de los errores estándar y los errores estándar robustos por grupos. De nuevo, con esta versión robusta del test, se rechaza la hipótesis de  $H_0: \gamma = 0$  por lo que concluimos que el modelo de Efectos Aleatorios no es el apropiado.

En este momento parece claro que el modelo más adecuado para nuestro objetivo es el de Efectos Fijos; pero también deberíamos responder a la cuestión de si es preferible un modelo global para todos los establecimientos y formatos o un modelo para cada establecimiento y formato.

La respuesta a esta pregunta está relacionada con una importante cuestión en marketing. Existe una sola política comercial o bien podemos hablar de una política

comercial diferente en función del formato y del tipo de establecimiento donde se comercializa el producto.

Para responderla utilizamos una versión distinta del test de agrupabilidad propuesto por Hsiao (2002, pág. 15 a 21). La hipótesis a plantear es:  $H_1: \beta_1 = \beta_2 = \dots \beta_N$ . Aceptar  $H_1$  significa que el modelo adecuado es  $y_{it} = \alpha_i + \beta' x_{it} + u_{it}$  mientras que rechazarla supone admitir que este modelo no es el adecuado y por tanto, tenemos un modelo para cada individuo.

Para este caso concreto los valores del estadístico son

$$903,734898 = F_1 \geq F_{(N-1)K, NT-N(K+1)}^{(\alpha)} = F_{(24-1) \cdot 36, 24 \cdot 108 - 24(36+1)}^{(5\%)} = F_{828, 1704}^{(5\%)} = 1,10254$$

por lo que se rechaza la hipótesis nula.

No existe un único modelo común a todos los individuos; que en términos más de marketing supone admitir que no existe una única política comercial para todos los clientes. Todavía resulta un poco prematuro saber qué tipo de política comercial existe o dicho de otra forma, que otros tipos de modelos pueden estar presentes.

A la vista de los resultados, parece más apropiado un modelo para cada uno de los formatos y establecimientos que un modelo más global. Antes acabar con esta cuestión se explora la inclusión de variables desfasadas de la variable a explicar cómo variables explicativas. Este planteamiento nos lleva a utilizar modelos dinámicos.

El modelo ahora es básicamente el mismo, la diferencia es que añadimos como variables explicativas la variable  $y$  desfasada hasta 4 veces. La razón es que la variable  $y$  tiene periodicidad semanal, de manera que es posible que en la variable influyan los valores que tomó la variable hace 1, 2, 3 y 4 semanas antes. Además y como afirman Hill, Griffiths y Lim, en Principles of Econometrics (2012; 365), la inclusión de variables desfasadas, ya sean endógenas o exógenas, ayuda a resolver problemas de autocorrelación presentes en el modelo. Esta es la línea de trabajo que se ha seguido.



La principal diferencia respecto a la estimación anterior se produce en el método de estimación, los estimadores que hemos utilizado anteriormente presentan problemas que es necesario resolver. La solución fue aportada por Nickell (1981). En el artículo Nickell demuestra que el estimador de Efectos Fijos está sesgado, y este sesgo no desaparece cuando el número de individuos del panel aumenta, por lo tanto el estimador de Efectos Fijos continúa siendo inconsistente para el caso de  $N$  y  $T$  pequeña. Sin embargo cuando  $T$  es grande, el estimador de Efectos Fijos resulta ser **consistente**. El estimador de Efectos Aleatorios resulta ser también sesgado en estos modelos dinámicos de datos de panel. Por lo tanto y gracias a la aportación de Nickell (1981) se utiliza el estimador de Efectos Fijos ya que  $T$  es bastante grande comparado con los tamaños habituales de otros paneles y  $N$  bastante pequeño.

Aunque los resultados son mejores que los del modelo anterior, el camino a seguir continúa siendo el mismo, ¿es preferible un modelo global para todos los establecimientos y formatos o un modelo para cada establecimiento y formato? Se trata de nuevo de contrastar la hipótesis  $H_1: \beta_1 = \beta_2 = \dots = \beta_N$ . Aceptar  $H_1$  significa que el modelo adecuado es  $y_{it} = \alpha_i + \beta' x_{it} + u_{it}$  mientras que rechazarla supone admitir que este modelo no es el adecuado y por tanto, tenemos un modelo para cada individuo. Los valores del estadístico obtenidos:

$$1900,508497 = F_1 \geq F_{(N-1)K, NT-N(K+1)}^{(\alpha)} = F_{(24-1) \bullet 40, 24 \bullet 104 - 24 \bullet (40+1)}^{(5\%)} = F_{920, 1512}^{(5\%)} = 1,10146$$

Nos permiten NO aceptar la hipótesis nula.

A la vista de los resultados, se pueden establecer las primeras conclusiones.

La primera, la construcción de un modelo que explique el comportamiento de la cuota de mercado en función de las variables de marketing, se ha demostrado como posible. Esto significa que podemos explicar el comportamiento de la cuota de mercado de una marca en función de las acciones comerciales propias y las de sus competidores.

La utilización de la técnica de los datos de panel, en cualquiera de los métodos de estimación utilizados, ha puesto de manifiesto que no existe un único modelo para

todos los formatos de la marca. Es decir, los  $\beta$  del modelo son distintos para cada uno de los individuos; por lo que no podemos hablar de un modelo único, sino de varios modelos. El tema es ¿cuántos modelos?, a esta pregunta se dará respuesta más adelante.

Desde el punto de vista de marketing, las conclusiones aparecen, lógicamente, desde otra óptica. El contraste de agrupabilidad utilizado puede traducirse como ¿existe una sola política comercial o bien hay una política comercial para cada formato/establecimiento? Con las conclusiones obtenidas del contraste, estamos en condiciones de afirmar que no existe una única política comercial, pero todavía es pronto para afirmar que existen tantas políticas comerciales como formatos y establecimientos existen. Para responder a esta pregunta se divide la muestra en función de los establecimientos y en función de los formatos, y de este modo contestar a la pregunta.

Se consideran ahora dos unidades de agrupación diferentes:

- Los establecimientos, por lo que nos encontraremos con cuatro paneles distintos formados por Hipermercados, Supermercados de 1000 a 2499 m, Supermercados de 400 a 999 m y Supermercados de 100 a 399 m.
- Los formatos, en este caso tendremos 6 paneles distintos, los formados por los formatos Hasta 350 g., De 350 a 499 g, De 500 a 899 g, De 900 a 1200 g y De 3000 a 4999 g.

Tanto en los establecimientos como en los formatos partimos de la hipótesis que los formatos/establecimientos que trabajan siguen el mismo modelo, esto es, los  $\beta$  asociados a ellos son los mismos dentro del establecimiento/formato.

Lógicamente serán diferentes a los de los otros establecimientos/formatos tal y como hemos demostrado indirectamente con los contrastes anteriores, ya que si bien se ha demostrado heterogeneidad entre los  $\beta$ , no se ha demostrado el origen de esa heterogeneidad; que bien puede partir de 3 fuentes distintas:

1. Heterogeneidad producida por el tipo de establecimiento.
2. Heterogeneidad producida por el formato de producto.
3. Heterogeneidad debida a la interacción tipo de establecimiento y formato de producto.

Esta última es la que se ha puesto de manifiesto en los contrastes de agrupabilidad anteriores, rechazando la existencia de un modelo único frente a un modelo específico para cada formato/establecimiento. A continuación se contrasta la existencia de la heterogeneidad de tipo 1 y 2.

La heterogeneidad producida por el tipo de establecimiento, desde la óptica del marketing implica contrastar la hipótesis de que existe un modelo para cada tipo de establecimiento frente a la no existencia de este modelo. Esta es una cuestión importante, como lo demuestra la existencia de figuras comerciales como el Key Account Manager en los Departamentos Comerciales de las empresas, que vislumbran la posibilidad de un tratamiento personalizado para cada tipo de establecimiento. Será por lo tanto esta hipótesis la que se pone a prueba.

En el segundo caso, para contrastar la heterogeneidad producida por el formato, desde el punto de vista del marketing significa dar un tratamiento homogeneizado al formato, independientemente del tipo de establecimiento donde se esté comercializando. Esta ha sido la práctica habitual en los departamentos de marketing, práctica que dada la evolución y concentración de la distribución se supone que ha debido ir cambiando con el paso del tiempo. Esta es la segunda hipótesis a contrastar.

Los resultados de los contrastes realizados nos muestran que:

- En los Supermercados de 100 a 399 m y en los Hipermercados se acepta la hipótesis nula, esto es, existe un solo modelo para todos y cada uno de los formatos que se comercializan en estos establecimientos.

Desde el punto de vista de marketing, significa aceptar que para los hipermercados, la figura del Key Account Manager está dando sus frutos,

realizando una comercialización del producto homogénea, sin distinción de formatos.

En el caso de los supermercados de 100 a 399 m, en los que también aceptamos la hipótesis, su interpretación en clave de marketing, es distinta, ya que la figura del Key Account Manager no tiene sentido en estos momentos. Parece más plausible pensar, que el elevado número de establecimientos con estas características, con un elevado número de propietarios, hace que para los fabricantes sea más sencillo tratarles a todos por igual que darles a cada uno un tratamiento específico como sucede en el caso de los hipermercados.

- En los Supermercados de 400 a 999 m y de 1000 a 2499 m. se rechaza la hipótesis nula, esto es, no hay un modelo único para todos, sino que parece haber un modelo para cada uno de los formatos.

Desde la óptica de marketing significa un tratamiento específico para cada formato. Esto parece apuntar a la idea de que en estos tipos de establecimientos es donde se está produciendo uno de los mayores cambios en la distribución española y que de momento resulta imposible un tratamiento homogéneo por formatos. Aun así esta conclusión será revisada un poco más adelante.

Cuando se repite el mismo contraste, pero esta vez la unidad de agrupación son los formatos y el panel estará formado por 4 individuos, los establecimientos donde se vende el producto. Los resultados permiten afirmar que:

- Existe un comportamiento homogéneo en el tratamiento de los formatos respecto a los establecimientos.
- Nos encontramos con 5 modelos distintos, uno para cada formato, en los formatos: Hasta 350 g., De 350 a 499 g., De 500 a 899 g., De 900 a 1199 g. y de 1200 a 1499 g. De tal manera que el comportamiento de los distintos tipos de establecimientos dentro del formato es homogéneo.

- El único formato donde se rechaza la hipótesis es en el formato de 3000 a 4999 g. La razón desde el punto de vista económico, que no estadístico, es que este formato tiene una presencia muy importante en los hipermercados, y son precisamente estos establecimientos los que tienen un comportamiento homogéneo como veíamos en el contraste anterior.

Tras estos contrastes podemos afirmar que el comportamiento de la empresa refleja dos dimensiones distintas pero a la vez que se superponen.

Una dimensión vertical, que proporciona un tratamiento homogéneo por formato independientemente del establecimiento y una dimensión horizontal con un tratamiento homogéneo por establecimiento y no por formato. Esta última dimensión resulta especialmente importante en los hipermercados.

Esta dimensión refleja el movimiento realizado por las compañías de gran consumo, dando un tratamiento personalizado a sus clientes, y este tratamiento ha comenzado por los hipermercados que son los que tienen un mayor poder para “exigirlo”.

En el futuro y con la concentración de la distribución que se está produciendo en España, especialmente en el supermercado grande y mediano, es posible que este tratamiento se extienda a otros tipos de establecimientos; cosa que hasta el momento no sucede.

En el capítulo 5 se aborda la construcción de los modelos para dar respuesta a las preguntas:

- ¿Qué variables son las que deben ser incluidas en cada uno de los modelos de datos de panel que hemos estimado?

- ¿Son las mismas variables en cada modelo o por el contrario cambian según los modelos?
- ¿La influencia de esas variables es idéntica en cada una de las unidades de panel consideradas?

La forma de proceder ha sido, estimar los modelos, eliminar las variables no significativas y quedarnos con los modelos más manejables y sencillos.

Esto se ha hecho para los 6 modelos considerados, uno para cada formato, menos el formato de 3000 a 4999 g., que ha sido sustituido por el modelo para hipermercados.

Cada modelo se estima y se eliminan las variables no significativas, resultando modelos más sencillos, en unos casos más que en otros. No todos los modelos presentan las mismas variables explicativas. Para cada modelo se realiza una interpretación de las variables que resultan significativas, tanto desde el punto de vista estadístico como del de marketing. Además se realiza un nuevo contraste para comprobar si variables influyen todas de la misma forma sobre la cuota en tienda. Aceptar la hipótesis nula supone reconocer que la influencia del tipo de establecimiento es la misma siempre. No aceptarla significa reconocer que alguna combinación de variable explicativa y tipo de establecimiento produce un efecto distinto (mayor o menor) sobre la variable a explicar. Por ejemplo, en el modelo planteado podría significar que la variable exposición especial tiene un efecto distinto en los Hipermercados que en los Supermercados Medianos.

Para cada modelo, se presentan cuadros y gráficos con los efectos de las variables según los tipos de establecimientos considerados. La única excepción es el modelo para hipermercados, donde los cuadros y gráficos hacen referencia a los formatos ya que el panel se ha definido de forma diferente.

El último paso ha sido comprobar que la forma funcional utilizada es la adecuada, mediante el contraste RESET que permite detectar variables omitidas y formas funcionales incorrectas. Cuando la forma funcional no ha resultado ser la correcta,

se ha vuelto a especificar buscando las no linealidades presentes y que no se habían incluido en el modelo original.

El capítulo 6 aborda las conclusiones, comienza recordando los dos objetivos principales. El primero, demostrar que es posible construir un modelo para explicar lo que sucede en el punto de venta, y no sólo de una forma descriptiva, sino de una forma interactiva, determinando qué variables influyen y cuáles no, y el peso de esta influencia.

El segundo objetivo era demostrar de una forma empírica que no existe una única política comercial para un producto en la empresa, sino que la política comercial es distinta según los formatos y los clientes.

El primer objetivo ha supuesto construir un modelo de marketing, tomando como variable a explicar la cuota de mercado y como variables explicativas, todas aquellas variables, tanto de la empresa como de la competencia, que puedan influir sobre ella.

Dentro de las diferentes metodologías se optó por construir un modelo con metodología de datos de panel. ¿Por qué este tipo de modelos? La razón de este tipo de modelos viene de considerar las circunstancias siguientes:

- Por un lado, con el tamaño de muestra con el que se trabaja, se hubiera podido estimar un modelo para cada formato y establecimiento. Este planteamiento nos hubiera llevado a estimar 24 modelos diferentes. Desde el punto de vista académico parece abordable, ahora bien, desde el punto de vista profesional, la dimensión del trabajo resulta un tanto exagerada, sobre todo si tenemos en cuenta que cada dos semanas sería necesario retroalimentar los modelos y actualizarlos. Ahora bien, y en un futuro como ampliación de investigación si cabe la posibilidad de trabajar en conjunto con todos los formatos utilizando la metodología de modelos SUR (regresiones aparentemente no relacionadas).

- En el otro extremo, se estimó un modelo único para todos los formatos y establecimientos. Este planteamiento nos hubiera llevado a demostrar que existe una única política comercial independientemente del formato comercializado o el punto de venta en el que se pone a disposición del consumidor. Sin embargo los contrastes de agrupabilidad que se realizaron echaron por tierra esta hipótesis, de tal manera que no podemos afirmar que exista una única política comercial.
- La última posibilidad que quedaba era buscar una metodología que permitiera agrupar, siempre que hubiera algún grado de homogeneidad, a los formatos y a los establecimientos. Esta metodología resultó ser los datos de panel, que imponiendo una restricción en la estimación (los parámetros son constantes entre los individuos del panel y las diferencias quedan recogidas en la constante y en la perturbación aleatoria), permitió la estimación de modelos por formatos y un modelo por tipo de establecimiento. Esta circunstancia está mucho más cerca del día a día de las empresas.

Respecto al segundo objetivo, una política comercial única por producto implica que no existe un tratamiento diferenciado para los distintos formatos que constituyen el producto. Nada más lejos de la realidad, los modelos planteados indican que las variables explicativas por formato son distintas. Esto quiere decir que la política comercial de la empresa varía según sea el formato e incluso el establecimiento, ya que los modelos también ponen de manifiesto que los pesos de las variables son distintos según sea el tipo de establecimiento considerado.

Realizadas estas consideraciones, en el último capítulo se estudia, en primer lugar, que variables son significativas dentro de los distintos modelos, independientemente del peso o influencia de cada una de ellas. Ordenadas según su frecuencia, las variables significativas en los distintos modelos son:

- **Retardo 1.** La historia más reciente resulta ser influyente en todos los modelos considerados, por tanto el pasado está condicionando el futuro, si bien es cierto que este condicionamiento sólo es de un periodo.



- ***X<sub>9</sub> Intensidad de la promoción de Precio del PRINCIPAL COMPETIDOR.***  
Esta es la variable que aparece con mayor número de influencias. Sólo en el formato de hasta 350 gr. no aparece, esto último es debido a las especiales características de este formato, ya que el dominio de la marca *Y* es aplastante, no teniendo prácticamente ninguna competencia.
- ***Retardo 4, X<sub>10</sub> Intensidad de la promoción de Precio del RESTO DE FABRICANTES y X<sub>13</sub> Cobertura de la Promoción sólo Exposición Especial de la Marca 1.*** En este epígrafe aparecen tres variables con el mismo número de menciones; el retardo 4, de nuevo la historia no tan reciente, acaba influyendo en el presente. La variable *X<sub>10</sub>* vuelve a poner de manifiesto la influencia de las acciones de precio de la competencia, agrupada como resto de fabricantes y por último la variable *x<sub>13</sub>*, esta variable es una variable de acción, una variable que es controlada por la marca, una variable que hace referencia a un marketing ofensivo, lo que hace la marca, frente a un marketing pasivo, lo que hace la competencia.
- ***X<sub>2</sub> Diferencial de precios respecto al Resto de Fabricantes, X<sub>7</sub> Distribución Ponderada Marcas de la Distribución, X<sub>17</sub> Cobertura Regalo Directo Marca 1.*** De nuevo aparece una variable relacionada con el tema precio: *X<sub>2</sub>*, y las restantes variables están relacionadas con otras variables distintas del marketing; *X<sub>7</sub>* hace referencia a la distribución del producto y *X<sub>17</sub>* a la influencia de un tipo especial de promoción como es el regalo directo. La primera y la última son variables que pueden ser manejadas por la empresa, mientras que la segunda, es una variable de la competencia y cuyo efecto se sufre.
- Por último aparecen un grupo de variables que sólo han sido significativas en cada uno de los modelos.

La conclusión resulta obvia; las variables que influyen a la hora de explicar la cuota de mercado de una marca son distintas según se trata del formato considerado. Si entendemos que un producto está compuesto por un conjunto de formatos, podemos afirmar que no existe un tratamiento único por producto sino un tratamiento

específico para cada formato. El conjunto de estos tratamientos forma lo que se llama la política comercial del producto.

Aún así, ha habido ocasiones donde el modelo lineal no ha pasado el contraste Reset de No linealidad, indicando que había que modificar la especificación del modelo, incluyendo algún componente no lineal que hasta ese momento no se había tenido en cuenta. Por lo tanto, se ha repetido este análisis con las nuevas especificaciones de la ecuación de regresión.

Después de realizarlo se obtienen tres conclusiones:

- Hay un grupo de modelos donde la forma funcional no presenta ningún tipo de No linealidad; estos son los modelos de Hasta 350 g e Hipermercados.
- Existe otro grupo de modelos donde la forma funcional presentaba problemas de No linealidad, y una vez corregida se ha obtenido una nueva especificación del modelo. Estos son los casos de los modelos de 350 a 499 g y de 1200 a 1499 g.
- Por último existe un modelo que presentando problemas de No linealidad, me ha resultado imposible encontrar la forma funcional correcta, y he preferido dejarlo con la ecuación original antes que complicarlos, en base al principio de parsimonia.

Por tanto estamos ante modelos distintos para los distintos formatos, reforzando aún más la idea de No existencia de una política comercial única para un producto.

La última parte del capítulo se centra en determinar qué variables son más influyentes en los distintos modelos. ¿Por qué es importante constatarla? Determinar que variables son más influyentes permite centrar los esfuerzos de las empresas en las variables que producen un mayor resultado en sus ventas.

Para medir esta influencia se utiliza un concepto muy habitual en la Economía: La Elasticidad. El camino a seguir ha sido calcular la elasticidad para cada una de las variables explicativas y posteriormente compararlas; las variables con una elasticidad más elevada, tendrán una influencia mayor sobre las ventas.

El siguiente paso ha sido definir claramente la elasticidad en un modelo lineal; es importante precisar el modelo de relación entre las variables porque su cálculo está influido por la relación existente entre las variables. Una vez definida, dentro de los dos tipos posibles de elasticidades a calcular, puntual y en media, se decidió calcular esta última.

El capítulo finaliza con el cálculo de la elasticidad en todos los modelos planteados y de esta forma puede determinarse, en cada modelo, qué variable explicativa tiene una mayor influencia sobre la cuota en tienda de la compañía que en la mayoría de los casos es la variable que tiene una mayor influencia sobre las ventas. El cálculo de la elasticidad ha permitido además estudiar la influencia de cada variable por tipo de establecimiento. Llegando a la conclusión que la elasticidad de una misma variable difiere de un tipo de establecimiento a otro. Es decir, que una reducción de precios del mismo importe, tiene una influencia diferente en un hipermercado que en un supermercado. Y cuando se trabaja con el modelo de hipermercados, se observa el mismo fenómeno, pero en este caso por formatos; una acción comercial en un formato produce resultados distintos que la misma acción en otro formato diferente.

## 2.- MODELOS DE MARKETING

El concepto de modelos está ampliamente difundido y utilizado en Economía y en multitud de otras disciplinas científicas. Por comenzar con una definición, podríamos decir que un modelo es una representación simplificada e idealizada de la realidad.

Sin embargo, el concepto de modelo en marketing es bastante más reciente. Wierenga (2008: 3) establece que el concepto de Modelos de Marketing aparece por primera vez a principios de los años 60; basándose en la publicación de tres libros sobre este tema: Bass et al. (1961), Frank et al. (1962) y Buzzel (1964). En estos libros se introduce el concepto de modelos de marketing, se discute sobre sus ventajas y se dan ejemplos de cómo los modelos de marketing pueden implementarse y utilizarse en las diferentes áreas del marketing: publicidad, planificación de medios, precios, etc.

Según Wierenga (2008:3), los factores que explican que aparezca el concepto precisamente en ese momento son:

- Los primeros ordenadores llegan a las empresas, posibilitando el tratamiento de grandes conjuntos de datos; es el comienzo de la era de la información.
- El Management se interesa cada vez más por las ciencias sociales, la estadística, las ciencias del comportamiento, etc.
- La Investigación de Operaciones o Management Científico que se había centrado hasta ese momento en los problemas de logística y transporte, diversifica su campo de atención abarcando prácticamente todas las áreas de la sociedad.

En estos 50 años, la evolución que han experimentado los modelos de decisión de marketing ha sido constante, tanto en las técnicas utilizadas como en las aproximaciones a distintas realidades empresariales donde aplicarlos.

Wierenga (2008:4) nos habla de que esta disciplina ha atravesado cinco etapas y las hace coincidir con las cinco décadas que van desde su aparición hasta el momento actual.

- **Los años 60: el comienzo.**

Leeflang et al (2000: 106) dice “esta etapa se caracteriza por la transposición directa de los métodos de la Investigación de Operaciones o del Management Científico al área de marketing”.

En esta etapa se aplican métodos de optimización, modelos de Markov, técnicas de simulación y teoría de juegos.

Quizás los modelos más relevantes de esta época fueron:

- El de Dorfman y Steiner (1954) que desarrolla un teorema a través del que se pueden determinar los valores de precio, publicidad y calidad que maximizan el beneficio.
- El de Vidale y Wolfe (1957). Los autores proponen un modelo para explicar las ventas como función de la inversión en publicidad y que tenga utilidad tanto descriptiva como predictiva. Como se verá más adelante, el modelo propuesto en esta tesis tiene relación con éste.
- El modelo de Bass (1969), con una gran repercusión en los libros de Dirección Comercial de nuestro país y que establece que para un bien duradero, la evolución de las ventas de su categoría dependen de la proporción de consumidores innovadores e imitadores.

- **Los años 70.**

Tal y como la califica Wierenga (2008:6) es la década dorada de los modelos de decisión de marketing. Su campo de aplicación crece exponencialmente, con una característica añadida, se desarrolla la identidad que en estos momentos tiene.

Tanto Wierenga (2008:6) como Leeflang et al (2000:107) coinciden; el punto de partida son los problemas de marketing, los modelos se vuelven más realistas. Se desarrollan como una disciplina a parte de la Investigación de Operaciones y así ha continuado desde entonces.

En esta época se produce una gran variedad de aproximaciones al marketing:

- Aparecen los primeros modelos estocásticos; especialmente los modelos de elección de marca por el consumidor, que más adelante recibirán una atención predominante. El modelo de Massy et al (1970) es un ejemplo de aplicación de métodos econométricos avanzados.
- La atención principal la recibieron los modelos que explicaban la relación existente entre un instrumento de marketing y las ventas. Por ejemplo la relación entre las ventas y los precios; o las ventas y la fuerza de ventas. Wierenga (2008:6) a este tipo de modelos les da el nombre específico de modelos de respuesta de marketing. Quizás uno de los más conocidos es el modelo de Kotler (1971). Pero tal y como dice Leeflang (1974) en una evaluación crítica de la literatura, estos modelos fallaban en la representación de la realidad, considerando que los efectos de la competencia eran estáticos y además sólo consideraban un instrumento de marketing, produciéndose un problema de omisión de variables.
- Se produce la aparición de los modelos con marca. Son modelos que comparten las características siguientes: tienen una estructura matemática específica, corren en ordenador, se usan para tomar decisiones de marketing y tienen un nombre o marca para distinguirlo de los demás. Los más conocidos fueron CALLPLAN de Lodish (1971), para la planificación de decisiones de llamadas de ventas, ADMOD de Aaker (1975) para la planificación de medios en publicidad y ASSESSOR de Silk y Urban (1978) para decisiones de nuevos productos.

- También en esta década aparece el concepto Sistema de Soporte de Decisiones de Marketing (MDSS) introducido por Little (1979b). El objetivo no era otro que el de reducir la distancia existente entre los modelos teóricos de marketing y la realidad que supone la toma de decisiones de marketing en la práctica.

- **La década de los 80.**

En esta década, tal y como reconoce Wierenga (2008:7) los modelos de respuesta de marketing han producido un número suficientemente elevado de estudios empíricos que es posible hacer generalizaciones. Es el momento de los Meta-análisis para diferentes instrumentos de marketing. Los estudios más citados de este tipo son el meta-análisis para publicidad de Asmus et al. (1984) y el de precios de Tellis (1988).

Como técnicas novedosas, es en este periodo donde aparecen los modelos de análisis conjunto (Wierenga 2008:7), técnica que continua utilizándose actualmente; y los modelos logit multinomiales (Leeflang et al 2000:108).

- **La década de los 90.**

Aunque no existe unanimidad entre Wierenga (2008) y Leeflang et al (2000), respecto al momento de su aparición, si es cierto que se produce una novedad importantísima en el punto de venta y que tendrá su efecto posterior en el marketing y por ende en la disciplina de los modelos de marketing. Esta novedad es la aparición de los scanner en los puntos de venta. El scanner posibilitó la recogida de un gran volumen de información y con una frecuencia hasta entonces desconocida. Dando lugar a lo que Blattberg et al (1994) llamó “la revolución de la información de marketing”.

Con tal volumen de datos a nuestra disposición era posible estudiar el efecto de las promociones sobre las ventas, o cualquier otra hipótesis que pudiera plantearse. Y no solamente eso, sino que las empresas

encargadas de recoger y tratar los datos proporcionados por los scanner; esto es, IRI y Nielsen, ven una oportunidad de negocio ya que existe una demanda por los directivos de marketing que necesitan conocer las causas de una caída o subida en las ventas. Es el momento de los modelos econométricos, tal y como dice Leeflang et al (2000:109), que estiman los efectos del marketing en las ventas. Tanto IRI como Nielsen desarrollan modelos parecidos, explotando sus datos y comercializándoles entre sus clientes; los directivos de marketing. Así IRI presenta PROMOTIONSCAN (Abraham y Lodish 1989) y Nielsen, SCAN\*PRO (Wittink et al. 1988). Ambos modelos intentan medir el efecto de las promociones sobre las ventas que se producen en el punto de venta. Son pues, desarrollos paralelos consecuentes con la estrategia competitiva que mantenían y siguen manteniendo ambas empresas.

Además en esta década comienzan a implantarse nuevas técnicas que vienen de otros campos. Es el momento de aparición de las Redes Neuronales; y de otras técnicas inductivas que extraigan el conocimiento de los datos. Esto es, aparece la semilla del concepto que más adelante fructificará como “data mining”.

Una cosa parece desprenderse de todo esto y que es sugerida por Leeflang et al (2000:109), el aumento de la disponibilidad de los datos de marketing, su adecuación y detalle supuso una fuerza importante para la implantación de los modelos.

- **Primera década del siglo XXI**

Según Wierenga (2008:8) el desarrollo más importante en estos años ha sido el cambio en la unidad de análisis, el epicentro de los estudios es el Consumidor Individual. Gracias al desarrollo de las tecnologías de la información, las empresas han sido capaces de recoger y almacenar un gran volumen de información de los clientes individuales. Estas bases de datos forman parte de lo que se llama sistema CRM ( Customer



Relationship Management). Los nuevos modelos de marketing desarrollados se centran en modelizar la adquisición y la retención de los consumidores, en predecir los consumidores que abandonarán la empresa y en ayudar a seleccionar los consumidores para una campaña de marketing específica. El concepto emergente es “el valor del consumidor individual” aglutinando estas propuestas en lo que se ha dado en llamar los modelos de valor del tiempo de vida del consumidor (CLV).

En fase todavía embrionaria, pero que seguramente su importancia crecerá exponencialmente en los próximos años, están los modelos de comercio electrónico.

Las dos opciones, CRM y comercio electrónico, han generado una cantidad ingente de datos, que producirá lo que Wierenga (2008:9) ha llamado una “segunda revolución en la información de marketing” y que lógicamente necesitarán de nuevos tipos de modelos para ser tratados.

## **2.1.- TIPOLOGÍAS DE MODELOS DE MARKETING**

Parece lógico que antes de profundizar en el tema de los modelos de marketing, deberíamos intentar clasificarlos para que su estudio resulte más sistemático y sencillo. La clasificación utilizada es la que propone Wierenga (2008:10), de tal forma que agruparemos a los modelos en las categorías siguientes:

### **2.1.1.- Modelos de Toma de Decisiones de los Consumidores.**

Se trata de modelos que intentan explicar cómo los consumidores tomamos nuestras decisiones de compra. La literatura que trata este tipo de modelos básicamente se centra dos tipologías:

- Modelos en los que se utiliza la metodología del análisis conjunto. Esta metodología permite medir las preferencias de los consumidores respecto a los atributos ofertados por un producto. Es especialmente adecuada cuando se pretende realizar una segmentación de mercados por ventajas buscadas, ya que identifica que atributos son los más valorados en cada producto y qué consumidores se sienten atraídos por dichos atributos y por lo tanto, por los productos. El análisis conjunto permite diseñar los productos de acuerdo a las preferencias de los consumidores.
- Modelos de Ayuda Interactiva a las Decisiones del Consumidor (ICDA). Tal y como dice Wierenga (2008:10) los consumidores actuales disponen de unas posibilidades de elección muy grandes y los ordenadores hacen posible pueden echarles una mano con ayudas interactivas en las decisiones. La metodología combina investigación del comportamiento del consumidor con modelos de elección e internet. Es una metodología novedosa sobre la que irán apareciendo aportaciones en los próximos años.

### **2.1.2.- Modelos de Marketing Mix.**

Son modelos que explican cómo influyen los instrumentos de marketing en las ventas.

El tema es interesante por dos motivos. El primero es que suponen un campo muy importante de los modelos de marketing desde que empiezan este tipo de modelos en los años 70. El segundo está relacionado con esta tesis, que como se verá, propone un modelo donde las variables explicativas son elementos del marketing mix.

Wierenga (2008: 11) clasifica los modelos según el instrumento de marketing utilizado para explicar las ventas, de esta forma habla de los siguientes tipos de modelos:

- Modelos de Publicidad, son modelos que explican cómo influye la publicidad en las ventas de la marca. Lógicamente estos modelos han ido evolucionando en el tiempo y así en los años 70, la principal preocupación era medir la influencia de la inversión en publicidad en las ventas o bien medir la influencia de un determinado medio en las ventas. Ejemplos de estas aplicaciones pueden encontrarse en Lilien et al (1992: 275-319). En estos momentos, están más interesados en explicar fenómenos como la exposición a los medios en Internet o el problema de la selección de canales.
- Modelos de Promoción de Ventas. En este caso lo interesante es modelizar y medir los efectos de la promoción de ventas.
- Modelos para Decisiones de Gestión de Ventas. En este caso se trata de explicar cómo las ventas responden ante cambios en la localización de recursos de ventas, en el diseño del territorio o en la estructura de la fuerza de ventas.

- Modelos de Respuesta Competitiva. Bajo este epígrafe el objetivo no es conocer la respuesta de las ventas a las variables de marketing mix de la empresa, sino la respuesta de las ventas a las variables de la competencia. En otras palabras, nos interesa saber, si una empresa competidora realiza una acción de precio; cómo afectará dicha acción a las ventas de nuestra empresa.

### **2.1.3.- Modelos de Marketing Centrados en el Cliente.**

Estos modelos se han desarrollado debido a la influencia que tiene el cliente individual en el marketing de hoy. Recordemos que la sociedad de la información e internet ha hecho posible recoger una gran cantidad de información sobre cada uno de los clientes individualmente. Se trata por tanto, de procesar toda esa información e intentar explicar el comportamiento de cada uno de los clientes.

Dentro de estos modelos podemos encontrar dos grandes grupos:

- Modelos del Valor del Cliente. El objetivo es determinar qué valor tiene para una empresa un cliente concreto. De esta manera la empresa puede calcular cuánto podría llegar a gastarse en ese cliente. Dentro de este contexto se ha desarrollado el concepto de “*Customer Lifetime Value*” que ha aparecido con mucha fuerza en marketing.
- Modelos de Decisión para la Gestión de Relaciones con Clientes (CRM). En este epígrafe se encuadran los modelos centrados en el proceso de captar, retener y desarrollar los clientes; o en la gestión de los clientes seleccionados.

#### **2.1.4.- Otras aproximaciones.**

Dentro de este epígrafe se encuentran contribuciones que todavía no han tenido un grado de aplicación o desarrollo amplio en el área de marketing. Así nos encontramos con modelos como:

- Modelos de Series Temporales en Marketing.
- Redes Neuronales y Algoritmos Genéticos en Marketing.
- Modelos desarrollados para algún sector específico (Cine, Industria Farmacéutica ...)
- Modelos donde se analizan los efectos financieros del marketing en las empresas.
- ...

Hecha esta descripción de los modelos de marketing, me centraré en los modelos de Marketing Mix y Series Temporales; ya que son los que más relacionados están con el modelo propuesto en esta tesis.

#### **2.2.- MODELOS RELACIONADOS CON EL MODELO QUE SE PRESENTA EN ESTA TESIS.**

El modelo desarrollado en esta tesis está relacionado con las tipologías de modelos expuestas anteriormente. Aun así, las tipologías sobre las que se realiza un zoom más detallado son:

- Modelos de publicidad.
- Modelos de promoción de ventas.
- Modelos de gestión de ventas.
- Modelos de respuesta competitiva.
- Modelos de series temporales en marketing.

Para cada tipo de modelo se estudia el estado de la cuestión de los modelos encuadrados en esa tipología y las semejanzas y diferencias con el modelo propuesto en la tesis.

¿Por qué estos modelos y no otros? Dos son los motivos de esta elección.

El primero es que se trata de modelos que estudian los efectos de variables de marketing mix sobre las ventas

Y el segundo es la aplicación de métodos habituales en la Econometría al área de Marketing.

Como se verá posteriormente estos son algunos de los puntos en común que tienen con el modelo que se expone en esta tesis.

### **2.2.1.- Modelos de Publicidad.**

El objetivo de los modelos de publicidad, como establece Danaher (Wierenga (2008: 81)) fue en un primer momento demostrar que la publicidad afectaba a las ventas o a la cuota de mercado, dada la importancia que tenía y tiene, la publicidad en el marketing mix. Logrado el objetivo, tal y como lo prueban los trabajos de Clarke (1976) y Leone y Schultz (1980), el interés cambia. Los trabajos siguientes se centran en construir modelos para optimizar el presupuesto de publicidad; son modelos de selección de medios y su interés recae en encontrar la combinación óptima de soportes publicitarios que maximicen la efectividad del presupuesto publicitario.

Por lo tanto, los modelos de publicidad han sido utilizados en tres áreas distintas:

- Modelos de Respuesta, que explican cómo influye la publicidad en las ventas o en la cuota de mercado de una empresa.

- Modelos de Planificación de Medios, que intentan maximizar la efectividad del presupuesto de publicidad, medido en GRP'S, sometido a una restricción presupuestaria. En estos modelos se supone dado el medio, y se trata de elegir la combinación óptima de soportes para un periodo determinado.
- Modelos de Selección de Medios, que intentan demostrar la efectividad de los distintos medios que pueden utilizarse en una campaña publicitaria.

De esta tipología establecida, los modelos que más se parecen al que se presenta en esta tesis son los Modelos de Respuesta. Siguiendo de nuevo a Danaher, (Wierenga (2008:83)), divide a estos modelos en cuatro tipos.

1. Modelos aislados. Son los modelos iniciales, tratan de explicar las ventas o la cuota de mercado en función de la inversión publicitaria, ignorando la existencia de otros instrumentos de marketing. Little (1979) revisa estos modelos obteniendo las conclusiones siguientes:
  - La función de respuesta de las ventas a la publicidad es cóncava o tiene aspecto de S. Esto significa que en un primer momento, una pequeña inversión en publicidad provoca una respuesta proporcionalmente más importante en las ventas, pero a medida que el mercado va madurando, son necesarias inversiones superiores en publicidad para mantener el resultado; pero llega un momento que un aumento de la inversión publicitaria no produce ningún efecto en las ventas. Esta conclusión ya está presente en un artículo anterior: Little (1970). Como se verá más adelante, este aspecto de la función de respuesta está presente en el modelo que aquí se plantea.
  - La publicidad realizada por los competidores también tiene impacto sobre las ventas de la empresa.

- La efectividad de la publicidad cambia en el tiempo; la respuesta de las ventas declina aunque la inversión permanezca constante.
- A estas conclusiones de Little, es necesario añadir otra más, de tipo econométrico. Y es que la omisión de otras variables explicativas, provoca lo que se conoce como el problema de sesgo por variables omitidas. De manera que es posible que la variable publicidad, esté teniendo en el modelo una influencia que no es la que efectivamente tiene en la realidad.

2. Modelos donde la publicidad es un elemento más del marketing mix. En este caso se mide la influencia que tiene la publicidad sobre las ventas, pero acompañada de otras variables del marketing mix como son: el precio, la promoción y la distribución. Uno de los modelos respuesta agregada más robustos es el de Wittink et al's (1988) que plantea un modelo SCAN\*PRO de datos agregados de establecimientos de un mercado

$$Ventas_{it} = \left( \prod_{j=1}^B (Precio_{jt})^{\beta_{ij}^{Precio}} \right) \left( \prod_{j=1}^B (Publicidad_{jt})^{\beta_{ij}^{Publicidad}} \right) \left( \prod_{k=1}^K (X_{ikt})^{\beta_{ik}} \right) \exp(u_{it})$$

Donde  $Ventas_{it}$  es la media geométrica de las ventas de todos los puntos de venta en el mercado para la marca  $i (i=1, \dots, B)$ ,  $Precio_{jt}$  es la media geométrica del precio en todos los establecimientos,  $Publicidad_{jt}$  es alguna medida del nivel de publicidad del mercado,  $(X_{i1t}, \dots, X_{iKt})$  son un conjunto de  $K$  covariables promocionales que hacen referencia al tipo de promoción y a la visualización de las acciones, y por último  $u_{it}$  es la perturbación aleatoria.

Este es el modelo base, a partir de él, se han ido realizando modificaciones en función de la aparición de paneles de datos (mejorar la regularidad y frecuencia de las mediciones), nuevas metodologías de análisis (logit multinomial) y tratamiento de problemas específicos (endogeneidad en el modelo y su solución con variables instrumentales, inclusión de variables retardadas y término de error autocorrelacionado)



3. Modelos de Interferencia de la Publicidad. Consideran a la publicidad realizada por la competencia como una interferencia en las ventas de la marca, ya que reduce su evolución. Estos modelos son interesantes, ya que como se verá más adelante, son los precursores de los modelos de respuesta de la competencia.
4. Modelos de Tracking Publicitario. Un problema tradicional al que se ha enfrentado el marketing, y por tanto la publicidad, es la relación entre ventas y publicidad. Una corriente muy extendida dice que las ventas son función de un conjunto de variables de marketing y no sólo de la publicidad, por lo tanto la responsabilidad de las ventas no puede recaer exclusivamente en la publicidad. Para resolver el problema surge el planteamiento siguiente: la publicidad genera conocimiento o notoriedad de la marca. Por tanto midamos el conocimiento de la marca generado por la publicidad y su relación con las ventas. Hollis (1995) demostró que las ventas y la notoriedad están altamente relacionadas.

### **2.2.2.- Modelos de Promoción de Ventas.**

Una promoción de ventas es una acción especial realizada por una marca o por un establecimiento que se produce en un lugar o momento determinado. Cuando se realiza una promoción se espera un impacto sobre las ventas. El estudio del impacto promocional en las ventas según Van Heerdle y Neslin (Wierenga 2008:123) depende del tipo de dato con el que contemos; así si trabajamos con datos de paneles de hogares los modelos desarrollados son absolutamente diferentes a los modelos desarrollados cuando se trabaja con datos de panel de establecimientos.

En el cuadro que se muestra a continuación, adaptado de Van Heerdle y Neslin (Wierenga 2008:124) se aprecian las diferencias entre estos dos tipos de modelos

## Paneles de Hogares

## Paneles de Establecimientos

Modelos de compra del producto	Modelos de ventas de la marca (Scan*Pro models)
Modelos de compra de la categoría	Modelos para medir los efectos de la aceleración y desaceleración de las ventas producidos por la promoción
Modelos de elección de la marca	
Modelos de cantidad comprada de la marca	Modelos de efectos cruzados entre marcas
Modelos que explican la aceleración o desaceleración que se produce en la compra como consecuencia de la promoción	

Los modelos más cercanos al que se presenta en esta tesis son los modelos de paneles de establecimientos ya que comparten el mismo tipo de dato; así que es sobre estos modelos sobre los que se realiza un análisis más exhaustivo.

El modelo más parecido, es un modelo que ya apareció en el punto anterior. Es un modelo del tipo Scan\*Pro desarrollado por Wittink et al (1988) y que recogen Van Heerde y Neslin (Wierenga 2008:124)

$$S_{bst} = \lambda_{bs} \mu_{bt} \prod_{b'=1}^B \left\{ PI_{b'st}^{\beta_{b'b}} \cdot \gamma_{1b'b}^{FEATONLY_{b'st}} \cdot \gamma_{2b'b}^{DISPONLY_{b'st}} \cdot \gamma_{3b'b}^{FEAT \& DISP_{b'st}} \right\} e^{u_{bst}}$$

Dónde:

$S_{bst}$	Ventas en unidades de la marca $b$ en el establecimiento $s$ en la semana $t$
$PI_{bst}$	Índice de precio (ratio entre el precio actual y el precio habitual) de la marca $b$ en el establecimiento $s$ en la semana $t$
$FEATONLY_{bst}$	Indicador de la promoción: 1 si hay promoción sin display para la marca $b$ en el establecimiento $s$ en la semana $t$ ; 0 en caso contrario.
$DISPONLY_{bst}$	Indicador del display: 1 si hay display sin promoción de la marca $b$ en el establecimiento $s$ en la semana $t$ , 0 en caso contrario.
$FEAT&DISP_{bst}$	Indicador de promoción & display: 1 si hay promoción y display de la marca $b$ en el establecimiento $s$ en la semana $t$ , 0 en caso contrario.

El modelo incluye un intercepto específico para la marca-establecimiento  $\lambda_{bs}$ , un intercepto específico para la marca-semana  $\mu_{bt}$ , elasticidad de precios propia ( $\beta_{bb}$ ) y cruzada ( $\beta_{b'b}, b' \neq b$ ) y multiplicadores propios ( $\gamma_{1bb}, \gamma_{2bb}, \gamma_{3bb}$ ) y cruzados ( $\gamma_{1b'b}, \gamma_{2b'b}, \gamma_{3b'b}, b' \neq b$ ) para la promoción, el display y la promoción & display

El modelo se linealiza tomando logaritmos y se estima por Mínimos Cuadrados Ordinarios.

Sus resultados muestran que la respuesta de las ventas ante un cambio porcentual en el precio tiene forma de S, esto es, hay un efecto umbral y saturación.

El índice de precio captura los efectos de los precios promocionales que pueden ser muy diferentes de los efectos de los precios habituales. Esta comparación entre los efectos de los precios promocionales y habituales fue estudiada por Mulhern y Leone (1991) y Bijmolt et al. (2005).

Además la construcción del modelo permite estudiar los efectos asimétricos entre marcas. Esto es, el impacto de una promoción en la marca A sobre las ventas de la marca B no tiene por qué ser igual al impacto de la promoción de la marca B sobre las ventas de la A. Esta característica es importante cuando se estudian las sustituciones o cambios entre marcas (brand switching).

Tomando como base este modelo van Heerde et al (2004), lo reformulan, su intención es analizar el efecto que las promociones tienen sobre las ventas de la marca. Entienden que una promoción produce:

- Un efecto sobre la marca.
- Un efecto cruzado entre marcas
- Un efecto entre periodos, desplazando ventas (adelantando o retardando) de un periodo a otro.
- Un efecto de expansión de la categoría, ya que la promoción produce un efecto llamada que atrae la atención sobre la marca que realiza la promoción y se extiende a toda la categoría de producto.

El modelo es una novedad ya que aunque el tema se había tratado con datos de hogares, hasta ese momento no se había abordado con datos de establecimientos y esto es realmente importante ya que son los datos de los que disponen los gestores de marketing.

Las variables explicativas son las mismas variables que en el modelo de Van Heerde y Neslin (Wierenga 2008:124), pero se añaden cuatro variables más ya que el precio índice se divide en cuatro categorías mutuamente excluyentes: Precio índice sin ningún tipo de acción, Precio índice sólo con promoción, Precio índice sólo con display y Precio índice con promoción y display simultáneamente.

La forma de proceder es utilizar estas variables explicativas, siempre las mismas, y correrlas sobre los cuatro efectos previamente descritos, para estimar la influencia de estas variables sobre los efectos de la promoción.

### 2.2.3.- Modelos de Gestión de Ventas.

Los modelos que aparecen en este epígrafe están relacionados con la gestión de la fuerza de ventas o con el territorio de ventas de una empresa. Una exposición bastante detallada de estos modelos puede verse en el capítulo escrito por Albers, S. y Mantrala, M. en Handbook of Marketing Decisions Models (Wierenga 2008:163).

Estos modelos se han desarrollado para explicar cómo un aumento de los esfuerzos de ventas: más vendedores, territorios más extensos, más llamadas telefónicas para realizar ventas, afectan a las ventas de la compañía. Desde la óptica de esta tesis, están más alejados que los anteriores modelos, pero tienen un cierto interés si consideramos los esfuerzos de ventas como una variable más del marketing mix.

Bajo esta premisa, el modelo a destacar es el modelo CALLPLAN de Lodish (1971), que ya habíamos mencionado anteriormente. Este modelo es el origen de los modelos de gestión de ventas y establece la forma funcional siguiente:

$$S_i = \alpha_i + (\beta_i - \alpha_i) \frac{h_i}{\gamma_i + h_i^{\delta_i}} \quad (i \in I)$$

Donde  $S_i$  representan las ventas,  $h_i$  son el número de llamadas y  $\alpha_i, \beta_i, \gamma_i, \delta_i$  son parámetros que tienen que estimarse. El interés del modelo recae en la forma funcional, que fue originalmente sugerida por Little (1970), Little llamó a esta función, función ADBUDG II y hace depender la cuota de mercado de inversión publicitaria.

$$Share = \min + (\max - \min) \frac{(adv)^\gamma}{\delta + (adv)^\gamma}$$

Posteriormente Lodish, adapta la forma funcional a sus variables objeto de estudio; esto es, las ventas como variable a explicar y las llamadas telefónicas (CALLPLAN) como variable explicativa.

Esta forma funcional es la que se utiliza en esta tesis, ya presenta varias ventajas:

- La variable cuota de mercado o la variable ventas del modelo de Lodish no pueden tomar nunca valores inferiores a cero.
- El máximo de la variable cuota de mercado no puede superar el valor 1. En el caso del modelo de Lodish, el máximo no se concreta en un valor determinado, pero si existe un máximo que viene determinado por la capacidad de producción de la compañía.

Los modelos posteriores al de Lodish han ido incluyendo otras nuevas variables, pero la forma funcional todavía está presente en muchos de ellos, aunque también combinada con funciones logit, que curiosamente presentan el mismo aspecto de S.

#### **2.2.4.- Modelos de Respuesta Competitiva.**

Los modelos de este tipo consideran que las acciones que realiza la competencia influyen, de una u otra forma, en las ventas de la empresa, por tanto, si deseamos modelizar las ventas o la cuota de mercado de la empresa, no sólo debemos incluir las variables de marketing de empresa, sino que tenemos que añadir las variables de marketing de la competencia.

Si seguimos a Leeflang (2008) podemos distinguir dos tipologías de modelos: así por un lado estarían los modelos de “Marca Individual”, que explican la demanda de una marca en función de variables de marketing mix propias y de la competencia. La evolución en estos modelos se ha producido, bien por sofisticar la función de demanda, incluyendo nuevas variables, o bien, por nuevas especificaciones de la función de demanda. El otro grupo de modelos llamados “Soluciones Simultáneas” se ha desarrollado de la mano de la Teoría de Juegos.

El modelo que se presenta en esta tesis está incluido dentro de los modelos de Marca Individual, por tanto vamos a centrar nuestra atención en este tipo de modelos, estudiando la evolución que han tenido en el tiempo.

- Modelo de Demanda Clásico

Uno de los primeros estudios en este campo es el Kotler (1965) que introduce la idea de la adaptación de la compañía ante cambios en la competencia. La especificación de un modelo como este podemos encontrarla en Leeflang (2008)

$$\hat{q} = \hat{a} + \hat{\beta}_p p + \hat{\beta}_{pc} p_c + \hat{\beta}_a a + \hat{\beta}_{ac} a_c$$

Donde  $\hat{q}$  es la demanda estimada de la marca  $j$  en unidades,  $a$  y  $a_c$  son las inversiones publicitarias de la marca  $j$  y de su principal competidor y  $p$  y  $p_c$  son los precios de la marca  $j$  y su principal competidor.

- Modelos LNB

Son una modificación del modelo de Kotler. Leeflang (2008) considera las funciones siguientes,  $Q$  para las ventas totales de la categoría de producto y  $m$  para la cuota de mercado de una marca.

$$Q = Q_T(p, a, k, p_c, a_c, k_c, ev) \text{ y } m = m_j(p, a, k, p_c, a_c, k_c)$$

donde  $Q_T$  y  $m_j$  son las formas funcionales para la Cantidad Total y la cuota de mercado de la marca  $j$ . Es destacable que la forma funcional está sin especificar, por tanto la aplicación práctica requerirá la elección de una forma funcional concreta y no genérica como se plantea en este modelo.

Las variables incluidas en el modelo son:

$p, p_c$  = los precios de la marca  $j$  y un índice de precios de la competencia.

$k, k_c$  = una medida de la calidad de la marca  $j$  y un índice de la medida de la calidad de la competencia.

$a, a_c$  = el gasto en publicidad de la marca  $j$  y un índice de los gastos en publicidad de la competencia.

$ev$  = un conjunto de variables ambientales que influyen en las ventas totales del producto.

Se trata de una aproximación sumamente teórica ya que incluye variables poco definidas (las variables ambientales) y otras variables difícilmente medibles (calidad de la marca), además de resumir la competencia en índices sin definir cómo calcularlos y sin tener en cuenta que la fusión de toda la competencia en un índice puede resultar contraproducente para la capacidad explicativa del modelo de cuota de mercado. Ahora bien, resulta muy interesante porque recoge los efectos de las acciones de la competencia en las ventas de la marca  $j$ .

En este mismo artículo, Leeflang (2008), nos muestra con un ejemplo cómo se pueden recoger las reacciones de los consumidores ante las acciones de la competencia. Así, la elasticidad de las ventas de  $j$ , respecto a su publicidad ( $\eta_{q,a}$ ) es igual a la elasticidad de la categoría de producto ( $\eta_{Q,a}$ ) más la elasticidad de la cuota de mercado respecto a la publicidad de  $j$  ( $\eta_{m,a}$ )

$$\eta_{q,a} = \eta_{Q,a} + \eta_{m,a}$$

Esta elasticidad captura el efecto de la publicidad de una marca en la demanda del consumidor, pero para capturar el impacto total debemos considerar cómo reaccionan los competidores a los cambios en la publicidad de la marca y cómo estas reacciones modifican la demanda del consumidor. Leeflang distingue entre *efectos parciales directos* y *efectos parciales indirectos* de la publicidad de la marca  $j$  sobre las ventas de la categoría de producto y sobre su propia cuota de mercado. Un efecto parcial indirecto se produce en el escenario siguiente: Si la marca  $j$  cambia su nivel de gasto publicitario, los competidores pueden reaccionar adaptando su propio nivel de gasto, lo que influirá tanto en la demanda total como en la cuota de mercado. En esta hipótesis se está asumiendo la teoría del oligopolio, la competencia reacciona con el mismo instrumento de marketing que ha producido su



reacción. Por tanto, un cambio en el nivel de publicidad produce un cambio en el nivel de publicidad de la competencia. Esta reacción es calificada por Leeflang como *reacción competitiva simple*. Una aproximación más realista, consistente y cercana al concepto de marketing mix, es la llamada *reacción competitiva múltiple*, según la cual, un competidor puede reaccionar ante un cambio en publicidad no sólo con un cambio en su publicidad sino también con un cambio en su precio y en cualquier otro instrumento de marketing.

Un caso de reacción múltiple competitiva es el siguiente:

$$\frac{\delta Q}{\delta a} = \frac{\delta Q_T}{\delta a} + \frac{\delta Q_T}{\delta p_c} \underbrace{\frac{\delta p_c}{\delta a}}_{\text{Efecto de la publicidad de la marca } j \text{ en el precio competencia}} + \frac{\delta Q_T}{\delta a_c} \frac{\delta a_c}{\delta a} + \frac{\delta Q_T}{\delta k_c} \frac{\delta k_c}{\delta a}$$

Efecto de la publicidad de la marca  $j$  en sus ventas
Efecto de la publicidad de la marca  $j$  en la Demanda Total
Efecto del precio competencia en la Demanda Total
Efecto de la publicidad de la marca  $j$  en el precio competencia

y

$$\frac{\delta m}{\delta a} = \frac{\delta m_j}{\delta a} + \frac{\delta m_j}{\delta p_c} \frac{\delta p_c}{\delta a} + \frac{\delta m_j}{\delta a_c} \frac{\delta a_c}{\delta a} + \frac{\delta m_j}{\delta k_c} \frac{\delta k_c}{\delta a}$$

Multiplicando la primera ecuación por  $\frac{a}{Q}$  y dentro del paréntesis multiplicamos y dividimos por cada instrumento, obtenemos la elasticidad para la categoría de producto.

$$\begin{aligned} \frac{a}{Q} \frac{\delta Q}{\delta a} &= \frac{a}{Q} \left( \frac{\delta Q_T}{\delta a} + \frac{\delta Q_T}{\delta p_c} \frac{\delta p_c}{\delta a} + \frac{\delta Q_T}{\delta a_c} \frac{\delta a_c}{\delta a} + \frac{\delta Q_T}{\delta k_c} \frac{\delta k_c}{\delta a} \right) \\ \frac{\delta Q}{\delta a} \frac{a}{Q} &= \frac{\delta Q_T}{\delta a} \frac{a}{Q} + \frac{\delta Q_T}{\delta p_c} \frac{\delta p_c}{\delta a} \frac{a}{Q} + \frac{\delta Q_T}{\delta a_c} \frac{\delta a_c}{\delta a} \frac{a}{Q} + \frac{\delta Q_T}{\delta k_c} \frac{\delta k_c}{\delta a} \frac{a}{Q} \\ \frac{\delta Q}{\delta a} \frac{a}{Q} &= \underbrace{\frac{\delta Q_T}{\delta a} \frac{a}{Q}}_{\eta_{Q,a}} + \underbrace{\frac{\delta Q_T}{\delta p_c} \frac{p_c}{Q}}_{\eta_{Q,p_c}} \underbrace{\frac{\delta p_c}{\delta a} \frac{a}{p_c}}_{\rho_{p_c,a}} + \underbrace{\frac{\delta Q_T}{\delta a_c} \frac{a_c}{Q}}_{\eta_{Q,a_c}} \underbrace{\frac{\delta a_c}{\delta a} \frac{a}{a_c}}_{\rho_{a_c,a}} + \underbrace{\frac{\delta Q_T}{\delta k_c} \frac{k_c}{Q}}_{\eta_{Q,k_c}} \underbrace{\frac{\delta k_c}{\delta a} \frac{a}{k_c}}_{\rho_{k_c,a}} \end{aligned}$$

Resumiendo:

$$\eta_{Q,a} = \eta_{Q,a} + (\eta_{Q,p_c})(\rho_{p_c,a}) + (\eta_{Q,a_c})(\rho_{a_c,a}) + (\eta_{Q,k_c})(\rho_{k_c,a})$$

Siendo:

$\eta_{Q_T,a}$  = elasticidad directa de la categoría de producto respecto a la publicidad de la marca  $j$ .

$\eta_{Q_T,u_c}$  = elasticidad de la categoría de producto respecto a los instrumentos de marketing de la competencia  $u_c = (p_c, a_c, k_c)$

$\rho_{u_c,a}$  = elasticidad de la reacción de los instrumentos de los competidores  $u_c = (p_c, a_c, k_c)$  con respecto a la publicidad de la marca  $j$ .

Este cálculo de la elasticidad se ha realizado para un cambio en el nivel de publicidad de la marca, pero puede generalizarse para cualquier otro instrumento de marketing.

Las condiciones de aplicabilidad de modelos de este tipo son enunciadas por Leeflang (2008), que establece que el modelo puede aplicarse si la empresa:

- No está muy interesada en los efectos de los competidores individuales y prefiere estudiar el efecto agregado del conjunto de marcas competidoras.
  - No se enfrenta a competencia vertical.
  - Especifica su marketing mix independientemente de los detallistas.
- Modelos LNB Extendidos con Funciones de Reacción de la Competencia Avanzadas.

El objetivo de estos modelos es estudiar qué variables del marketing mix provocan una variación en los precios (que es el caso que se presenta) o en cualquiera otro instrumento del marketing mix.

Modelos como estos se construyen cuando se relajan algunas de las hipótesis de partida de los modelos anteriores, más concretamente:

- Tiene en cuenta a todas las marcas separadamente y no en un agregado de marcas como se hacía en el modelo anterior. De manera que  $x_{ijt}$  es el valor del instrumento  $i$ , para la marca  $j$ , en el momento  $t$ .
- El modelo anterior consideraba que las decisiones de marketing mix se tomaban independientemente de los detallistas. Esta hipótesis es demasiado simplificadora de la realidad comercial. Leeflang y Wittink (1992) establecen cuatro tipos de reacciones de la competencia:
  - Movimientos paralelos. Se producen cuando hay fluctuaciones paralelas entre el precio o las promociones de la marca  $i$  y el precio y las promociones de la marca  $j$  en el mismo periodo  $t$  o con un periodo de retraso.
  - Reacciones dominadas por el detallista: corto plazo. Se observa la categoría de producto, si las actividades promocionales son frecuentes, una marca, seguida de otra marca en la semana siguiente o con una semana de retraso. El aumento en ventas de una marca en un periodo se ve compensado por el aumento de otra marca en el periodo siguiente.
  - Reacciones dominadas por el detallista: largo plazo. Los detallistas cambian el precio o realizan promociones en la marca  $j$  como consecuencia de los cambios realizados por otros detallista en variables similares de la misma marca o marcas competidoras. Este tipo de reacción no es posible estudiarla en el modelo que se presenta, ya que los datos están agrupados por detallista y resulta imposible observar el comportamiento individualizado de cada uno de ellos.
  - Reacciones dominadas por el fabricante. Esta reacción se observa sólo si el detallista coopera con el fabricante. Esta

cooperación resulta de plan promocional anual. En el modelo de la tesis tampoco es posible observar este tipo de reacción ya que aunque puede observarse la acción del fabricante, esta acción no viene desglosada por detallista.

Un ejemplo de modelos de este tipo pueden encontrarse en Leeflang y Wittink (1992) en el que estudian la influencia de los instrumentos de marketing siguientes: precio ( $p$ ), reparto de muestras ( $sa$ ), reembolso ( $rf$ ), oferta de bonos ( $bo$ ) y publicidad del detallista ( $ft$ ).

El modelo planteado es:

$$\ln\left(\frac{p_{jt}}{p_{jt-1}}\right) = \alpha_j + \sum_{r=1, r \neq j}^n \sum_{t^*=1}^{T^*=1} \beta_{jrt^*} \ln\left(\frac{p_{r,t-t^*+1}}{p_{r,t-t^*}}\right) + \sum_{t^*=2}^{T^*=1} \beta_{jft^*} \ln\left(\frac{p_{j,t-t^*+1}}{p_{j,t-t^*}}\right) \\ + \sum_{r=1}^n \sum_{t^*=1}^{T^*=1} \sum_{x=1}^4 \tau_{xjrt^*} (x_{r,t-t^*+1} - x_{r,t-t^*}) + \varepsilon_{jt} \\ \text{para } j = 1, \dots, n \text{ y } t = T^*+2, \dots, T,$$

$x = 1 = sa$ ; muestras

$x = 2 = rf$ ; reembolso

$x = 3 = bo$ ; bonos

$x = 4 = ft$ ; publicidad detallista

Dónde:

$T^* =$  máximo número de retardos ( $T^* = 10$ )

$T =$  número de observaciones disponibles

$n =$  número de marcas

$\varepsilon_{jt} =$  término de perturbación aleatoria

En el modelo planteado se estudia la variación en el precio de la marca  $j$ , respecto al periodo anterior, como consecuencia de las variaciones en el precio de otras marcas, de las variaciones en el pasado de la propia marca y de cambios producidos en los distintos instrumentos de marketing utilizados. Es por este motivo que se habla de reacción de la competencia.

- Modelos LNB Extendidos con Funciones de Demanda Avanzadas.

En estos modelos, desarrollados por Leeflang y Wittink (1996), utilizan como variable a explicar el cambio producido en la cuota de mercado. Y ese cambio puede deberse tanto a una modificación en el precio de la propia marca, de las marcas competidoras o cambios en las restantes variables del marketing mix, tanto de la marca como de la competencia.

$$\underbrace{\ln\left(\frac{m_{jt}}{m_{jt-1}}\right)}_{\text{Tasa de variación en la cuota de mercado respecto al periodo anterior.}} = \lambda_j + \sum_{r=1}^n \sum_{t^*=1}^{T^*=1} \gamma_{jrt^*} \underbrace{\ln\left(\frac{p_{r,t-t^*+1}}{p_{r,t-t^*}}\right)}_{\text{Variación en el precio marca } r \text{ respecto al periodo anterior}} + \sum_{r=1}^n \sum_{t^*=1}^{T^*=1} \sum_{x=1}^4 \xi_{xjrt^*} \underbrace{\left(x_{r,t-t^*+1} - x_{r,t-t^*}\right)}_{\text{cambio en las variables de marketing de las marcas}} + u_{jt}$$

El modelo tiene en cuenta varias marcas ( $r=1,\dots,n$ ) y varios periodos ( $t^*=1,\dots,T^*+1$ ).

Tiene la ventaja que puede estimarse para todas las marcas presentes en el mercado, de manera que puede establecerse un diagnóstico de que instrumentos de marketing resultan más efectivos a la hora de explicar la variación en la cuota de mercado.

Como se verá más adelante, este modelo guarda cierta similitud con el modelo planteado en esta tesis.

- Modelos de Series Temporales, Modelos VARX.

Son modelos que utilizan metodología de series temporales; los modelos Vector Autorregresivos con variables exógenas (VARX) y que se aplican en entornos de respuesta competitiva, concretamente se pueden aplicarse si:

- Deseamos tener en cuenta los efectos dinámicos que tienen los instrumentos de marketing sobre las ventas de una marca individual en el mercado.
- Distinguir entre efectos (de los instrumentos de marketing) inmediatos, brutos: son los efectos directos e indirectos de los instrumentos de marketing y netos: son los efectos directos e indirectos una vez que se tienen en cuenta las reacciones de la competencia.

En la aplicación de modelos VARX encontramos modelos de Funciones de Respuesta y de Funciones de Reacción, tal y como hemos visto previamente, los primeros estudian la influencia sobre las ventas que tienen los instrumentos de marketing, los segundos estudian la influencia que tienen los instrumentos de marketing en las reacciones (cambios) en las variables de marketing (especialmente cambios en precios) de las marcas. En este caso nos centraremos en los modelos de Funciones de Respuesta por ser más parecidos al modelo que se presenta aquí.

El modelo que estudiaremos es un modelo propuesto por Leeflang (2008) que tal y como él mismo afirma es una adaptación de un modelo SCAN\*PRO de AC Nielsen propuesto por Christen et al (1997).

$$\ln S_{qi,t} = \alpha_{qi} + \sum_{k=1}^2 \sum_{j=1}^n \sum_{t^*=0}^{P_{ijk}^{SP}} \beta_{PIjk,t^*} \ln PI_{qjk,t-t^*} + \sum_{j=1}^n \sum_{t^*=1}^{P_{ij}^{SS}} \phi_{ij,t^*} \ln S_{qj,t-t^*} \\ + \sum_{j=1}^n \beta_{Fij} F_{qj,t} + \sum_{j=1}^n \beta_{Dij} D_{qj,t} + \sum_{j=1}^n \beta_{FDij} FD_{qj,t} + \varepsilon_{qi,t} \\ q = 1, \dots, Q; \quad i = 1, \dots, I; \quad \text{y} \quad t = 1, \dots, T$$

dónde:

$\ln S_{qi,t}$  = logaritmo de las ventas de la marca  $i$  en el almacén  $q$  en la semana  $t$ ,  
 $\ln PI_{qjk,t}$  = logaritmo del índice de precios (actual respecto al habitual) de la marca  $i$   
 en el almacén  $q$  en la semana  $t$  ( $k = 1$  denota la característica de promoción / display  
 con la reducción de precios, y  $k = 2$  denota reducción de precio exclusivamente),  
 $F_{qj,t}$  = Variable ficticia que indica la existencia de una promoción "no de precios" de la  
 marca  $j$  en el almacén  $q$  en el periodo  $t$ ,  
 $D_{qj,t}$  = Variable ficticia que indica la existencia de una exposición especial en una promoción  
 "no de precio" de la marca  $j$  en el almacén  $q$  en el periodo  $t$ ,  
 $FD_{qj,t}$  = Uso combinado de promoción y exposición especial de una promoción "no de precio"  
 de la marca  $j$  en el almacén  $q$  en el periodo  $t$ ,  
 $\alpha_{qi}$  = Punto de corte con el eje de las  $Y$  de la marca  $i$  en el almacén  $q$ ,  
 $\beta_{PIijk,t^*}$  = Elasticidad de ventas (agrupada) de la marca  $i$  respecto al índice de precio de la marca  $j$   
 $\varphi_{ij,t^*}$  = Elasticidad (agrupada) de sustitución de las ventas de la marca  $i$  respecto a las  
 las ventas de la competencia ( $j$ ) en la semana  $t$  ( $i \neq j$ ),  
 $\beta_{Fij}, \beta_{Dij}, \beta_{FDij}$  = Efectos de la promoción "sólo" ( $F$ ), exposición "sólo" ( $D$ ), y promoción y exposición ( $FD$ ),  
 $PI_{ijk}^{SP}$  = Número de retardos  $k$  de la variable índice de precio de la marca  $i$  incluida en la ecuación  
 de la marca  $j$ ,  
 $P_{ijk}^{SP}$  = Número de retardos de la variable ventas de la marca  $i$  incluida en la ecuación de la marca  $j$ ,  
 $n$  = Número de marcas en la categoría de producto,  
 $Q$  = Número de almacenes, y  
 $\varepsilon_{qi,t}$  = Perturbación aleatoria

El modelo incluye diferentes variable promocionales, tanto propias como de la competencia: índice de precio, promoción sólo, exposición sólo, y promoción y exposición conjuntamente. Además Leeflang (2008) extiende el modelo incluyendo, y esta es la novedad, los efectos dinámicos de las promociones de precio (respuesta retardada) y reforzamiento de los efectos de compra (mediante las ventas retardadas).

Incluye dos tipos de variable de promoción de precios:  $K=1$  cuando la promoción de precios está anunciada y  $K=2$ , cuando la promoción de precios no está anunciada.

El modelo explica las ventas de una marca determinada, en un establecimiento concreto y en un momento determinado. Por tanto los datos de partida tienen que estar referidos a establecimientos determinados ( $q$ ), este tipo de análisis sólo es posible realizarlo contando con la colaboración, como es el caso, de AC Nielsen o el de su competidor (IRI), ya que estos

datos primarios no están disponibles para las empresas clientes o para los establecimientos colaboradores.

La técnica resulta adecuada ya que trata conjuntamente, la ventas de los  $q$  establecimientos, en función de las variables de marketing. Pero plantea un inconveniente que el propio Leeflang (2008) reconoce. Cuando se pretende testar la igualdad de las pendientes entre los establecimientos ( $q$ ) la hipótesis resulta rechazada, esto significa que no hay un modelo único, sino un modelo para las ventas de la marca  $i$  en cada establecimiento. Esto implica que si  $q=500$  establecimientos, tendríamos 500 modelos de ventas de la marca  $i$  en función de las variables de marketing. La influencia que tienen esas variables de marketing sobre las ventas, son diferentes ya que hemos rechazado la hipótesis de que su influencia era la misma. Desde el punto de vista de la aplicabilidad, nos encontraríamos con un problema de manejo del modelo por la complicación que supone su mantenimiento.

### **2.3.- ¿QUÉ TIENE EN COMÚN Y EN QUÉ SE DIFERENCIA EL MODELO QUE SE PRESENTA EN ESTA TESIS CON MODELOS ANALIZADOS?**

En este apartado pasaremos revista a las analogías y diferencias existentes entre estos modelos y el modelo que se presenta en la tesis.

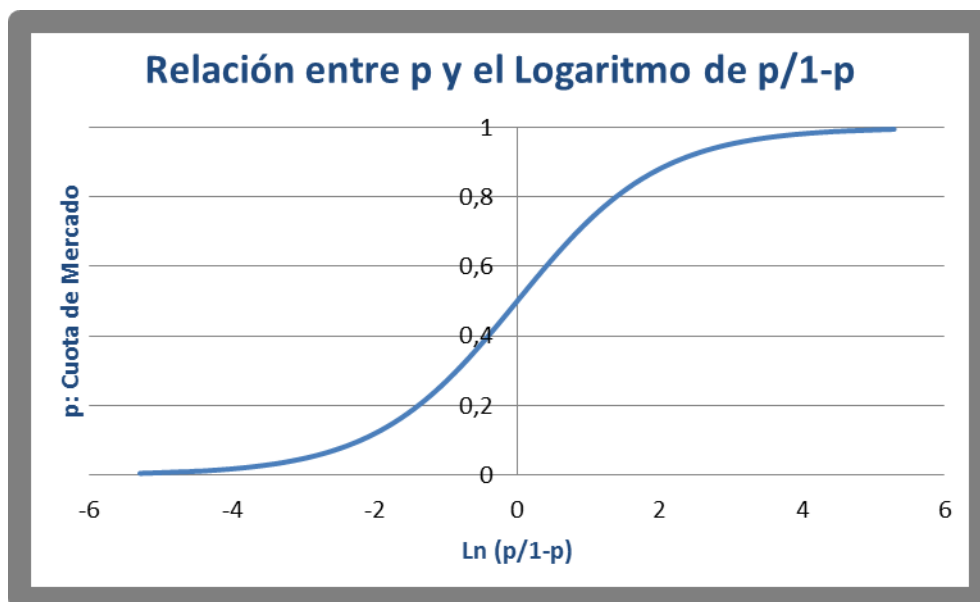
Siguiendo el orden de presentación de los modelos, la primera analogía se produce respecto a los modelos de publicidad. En estos modelos, resultado de la preocupación de sus autores, la función de respuesta utilizada para especificar la relación entre las ventas y la publicidad, es cóncava o tiene forma de S.

Esta forma funcional tiene sentido, ya que es esperable que en un primer momento, un ligero aumento de la publicidad produzca un aumento notable de las ventas, pero cuando el mercado se va saturando, el aumento de la publicidad se ve correspondido con ligeros aumentos de las ventas. Además si seguimos a Leeflang y Wittink (2000), que en su artículo recogen una idea de Little respecto a la estructura de los modelos. La idea no es otra que los modelos deben ser simples,



completos (incluir las características más importantes), adaptativos y robustos. . La robustez puede lograrse con una estructura que restrinja las respuestas a un rango de valores que tenga sentido. Esto es, si los valores de la variable están restringidos, como por ejemplo una cuota de mercado que estará comprendida entre 0 y 1, el modelo debería satisfacer las mismas restricciones. En este caso, estos modelos serían calificados como consistentes.

El modelo que se presenta tiene como variable a explicar la variable  $\ln\left(\frac{p}{1-p}\right)$  donde  $p$  es la cuota de mercado en tienda. La variable muestra la relación existente entre la cuota de mercado de la marca  $p$ , frente a la cuota de mercado de la competencia  $1-p$ , para posteriormente calcular el logaritmo neperiano de esta relación. La ventaja de esta variable es que el aspecto en S buscado por los modelos publicitarios está garantizado y cumple por tanto con las restricciones impuestas por la realidad, ninguna cuota de mercado puede ser negativa o superior a 1.



Añade una ventaja adicional y es que pueden introducirse todas las variables explicativas en una función lineal, por lo que la estimación será siempre más sencilla que en otros tipos de formas no lineales.

La segunda analogía se produce respecto al tipo de modelo. El modelo planteado en la tesis es un modelo tipo SCAN\*PRO. Y es un modelo SCAN\*PRO porque utiliza datos de panel de establecimientos proporcionados en este caso por Nielsen. Ejemplos de este tipo de modelos los encontramos en el epígrafe de publicidad, concretamente el modelo de Wittink et al's (1988)

$$Ventas_{it} = \left( \prod_{j=1}^B (Precio_{jt})^{\beta_{ij}^{Precio}} \right) \left( \prod_{j=1}^B (Publicidad_{jt})^{\beta_{ij}^{Publicidad}} \right) \left( \prod_{k=1}^K (X_{ikt})^{\beta_{ik}} \right) \exp(u_{it})$$

Y en el modelo de Van Heerde y Neslin (Wierenga 2008:124) en el epígrafe de modelos de promoción de ventas

$$S_{bst} = \lambda_{bs} \mu_{bt} \prod_{b'=1}^B \left\{ PI_{b'st}^{\beta_{b'b}} \cdot \gamma_{1b'b}^{FEATONLY_{b'st}} \cdot \gamma_{2b'b}^{DISPONLY_{b'st}} \cdot \gamma_{3b'b}^{FEAT\&DISP_{b'st}} \right\} e^{u_{bst}}$$

Ambos son modelos SCAN\*PRO, el primero mide la influencia del precio y la publicidad en las ventas de la marca en los establecimientos, y el segundo la influencia del precio, la existencia de promoción (no indica tipo de promoción), la existencia de display y la existencia conjunta de promoción y display sobre las ventas de la marca en los puntos de venta. Estos modelos junto con el que se presenta aquí, comparten el tipo de dato, son datos que muestran las ventas de las marcas en los distintos establecimientos, durante un periodo determinado. Sin embargo las diferencias son notables respecto a la forma funcional, el tipo de variables incluidas en los modelos, el modelo presentado es más prolijo y recoge con mayor detalle el tipo de promoción realizado por la marca y la intensidad de la promoción (importancia de los establecimientos que han participado en la promoción). Respecto a la variable “Display”, entendida como la forma de anunciar la promoción, el modelo la recoge dividiéndola en dos variables diferentes, para analizar su influencia: Folleto y Cabecera de góndola, permitiendo su análisis por separado y de forma conjunta (interacción), como aparece en el modelo de Van

Heerde y Neslin. El precio es tratado con mayor detalle, no se construye un índice único como aparece en estos modelos, sino diferentes índices para cada marca competidora.

La tercera analogía se produce con los modelos de Gestión de Ventas. Estos modelos utilizan una forma funcional que garantiza que la variable a explicar tenga forma de S, la aproximación de Little (1970) introdujo como variable a explicar la cuota de mercado y además introducen como variables explicativas otras variables distintas a la publicidad y la promoción. El modelo objeto de esta tesis presenta la misma forma funcional, con la cuota de mercado, pero simplificando su estimación, ya que veremos que se estima linealmente e introduce variables relacionadas con la distribución del producto, por lo tanto, variables no exclusivamente publicitarias o promocionales como hacen estos modelos de Gestión de Ventas.

Respecto a los modelos de Respuesta Competitiva, la analogía se produce porque ambos modelos introducen como variables explicativas, variables que recogen las acciones de marketing realizadas por la competencia. Se entiende que la cuota de mercado de una marca es fruto de las acciones de la propia marca y de las acciones de sus competidores, si esto es así, es lógico incluir esas acciones de la competencia en el modelo. Como se comprobará, esas acciones de la competencia resultan ser estadística y económicamente significativas, reforzando por tanto la hipótesis de su inclusión.

Por último, respecto a los modelos de series temporales, las analogías se producen en que ambos incluyen variables promocionales tanto propias como de la competencia. Otra similitud importante está en la utilización de variables retardadas, se trata en ambos casos de modelos dinámicos. Las grandes diferencias están:

- En el tipo de dato que utilizan; mientras que los modelos de serie temporal, trabajan con datos desagregados a nivel de establecimiento, el modelo objeto de esta tesis trabaja con datos agregados (globalmente, por formato o por tipo de establecimiento). Las empresas sólo pueden acceder este tipo de información, la información desagregada es posible cuando se trabaja con los datos primarios procedentes de la colaboración de Nielsen o Iri.

- El método de estimación, los modelos de series temporales permiten estimar un modelo para cada establecimiento, pero tal y como Leeflang (2008) reconoce, no es posible aceptar la hipótesis de la igualdad de las pendientes entre los establecimientos. En otras palabras, no hay un modelo único, sino un modelo para cada establecimiento, por tanto cuando se trata de aplicar y mantener el modelo en la práctica, resulta muy complejo. Si trabajamos con 1000 establecimientos tendremos 1000 modelos diferentes. En el modelo planteado en la tesis, se utiliza el método de estimación con datos de panel; como se verá, el método permite rechazar la hipótesis de un modelo único, como el caso anterior, pero sin embargo si permite una agrupación, ya sea por tipo de establecimiento, ya sea por formato. En la práctica esto se traduce en modelos más sencillos de mantener y de aplicar a situaciones concretas empresariales.

### 3.- ¿POR QUÉ DATOS DE PANEL?

Un panel es, siguiendo la definición que nos da Yustas (2014: 13) una muestra representativa de la población que permanece constante en el tiempo y que se estudia de forma repetitiva y periódica a intervalos regulares de tiempo. Es, por lo tanto, un instrumento de investigación sumamente útil siempre que deseemos estudiar la evolución en el tiempo de las unidades muestrales.

La metodología de análisis viene determinada, en muchos casos, por el tipo de dato disponible. Los datos utilizados en esta tesis han sido proporcionados por Nielsen; y son los datos habituales que esta compañía entrega a sus clientes. Son datos agregados que muestran la evolución en el tiempo de diferentes magnitudes de marketing. Esta agregación puede realizarse por formato, por tipo de establecimiento e incluso por zona de ventas.

El nivel de agregación de los datos hace imposible utilizar metodologías como las expuestas en modelos de series temporales del capítulo anterior. Este tipo de modelos sólo es posible ponerlo en práctica cuando se tiene información para cada uno de los establecimientos.

Además, la aplicación de esa metodología supone el problema determinar si existe un modelo para cada establecimiento o bien un modelo único y global. Estamos ante los dos extremos y nunca este tipo de soluciones son las más adecuadas, el primero, un modelo único supone admitir que la realidad comercial es idéntica independientemente del formato que se comercialice o el establecimiento donde se venda, realidad demasiado simplificadora y que será necesario contrastar. El otro extremo supone admitir un modelo diferente para cada formato y establecimiento, realidad compleja, pero que desde el punto de vista de gestión de las empresas implica una dificultad tremenda cuando se trata de analizar la realidad comercial. Y un problema de mantenimiento de los modelos cuando una vez estimados se trata de explotarlos regularmente.

La pregunta por tanto es obvia, ¿existe algún punto intermedio entre ambas hipótesis?, ¿existe alguna técnica que nos permita encontrar puntos en común entre

formatos y establecimientos y haga posible formular modelos más adaptados a las diferentes realidades comerciales pero sin estar subyugados por la necesidad de un modelo para cada establecimiento?

Es aquí donde los modelos de datos de panel se han mostrado especialmente útiles, proporcionando información respecto a la existencia de un único modelo, que como se verá en capítulos posteriores de la tesis, resulta ser una hipótesis imposible de mantener. O bien, mostrando el camino de una agrupación de los modelos, que sin llegar a ser un modelo para cada establecimiento o producto si puede encontrarse un punto en común que permita analizar la heterogeneidad de la realidad comercial y el compromiso con la aplicación y mantenimiento de estos modelos.

En lo que sigue, veremos una revisión de los distintos métodos de estimación de datos de panel que existen y que muchos de ellos se han utilizado en esta tesis

### **3.1.- EL MODELO LINEAL DE DATOS DE PANEL**

Los datos de panel son un conjunto de datos ordenados en el tiempo para los mismos individuos, siendo relevante la ordenación temporal de los datos y no la de los individuos. Por lo tanto, si modificamos la ordenación de los individuos pero dejamos invariable la ordenación temporal, los resultados serán exactamente los mismos.

La gran ventaja que presentan los datos de panel es que es posible observar las relaciones dinámicas que se producen en los individuos, cuestión que resulta imposible de determinar cuando trabajamos con datos de sección cruzada.

La forma habitual en la que se presentan los modelos de datos de panel es:

$$y_{it} = x_{it}\beta + u_{it} \quad i = 1, \dots, N, \quad t = 1, \dots, T$$

El subíndice  $i$  denota al individuo y el subíndice  $t$  denota el momento tiempo al que está referida la observación del individuo. El parámetro no está acompañado de ningún subíndice, quiere esto decir que permanece constante en el tiempo y para todos los individuos, por lo que las diferencias entre individuos se recogen en el intercepto ya que por construcción los parámetros son los mismos para todos los individuos.

Como en todo modelo econométrico el interés reside en estimar  $\beta$  y de esta manera, cuantificar la relación existente entre las variables  $X$  y la variable  $Y$ .

La ventaja o el inconveniente, es que dentro de los modelos de datos de panel tenemos a nuestra disposición, distintos métodos de estimación del parámetro, y que son en los que nos detendremos en este capítulo.

### **3.1.1.- MÉTODO MCO AGRUPADO.**

Este método no tiene en cuenta la información que proporcionan las diferencias entre individuos, utilizando exclusivamente su evolución en el tiempo. Por este motivo el modelo puede escribirse de la manera siguiente:

$$y_t = x_t\beta + u_t \quad t = 1, \dots, T$$

Es de notar que desaparece el subíndice  $i$ , las variables explicativas son tratadas como un todo, muy variado ya que recoge información de muchos individuos en el tiempo, pero desprecia la información de a que individuo pertenece cada una de las variables. Como se verá más adelante, cuando se estima un modelo con este método aparece un solo intercepto.

Es el método más cercano al habitual MCO.

Este modelo puede parecer sumamente restringido ya que  $\beta$  es el mismo en cada periodo de tiempo y para cada individuo; sin embargo si introducimos variables

ficticias temporales dentro de las  $x_t$  podemos conseguir que los parámetros cambien en el tiempo.

Algo más a añadir; cualquier cosa que podamos hacer con datos de sección cruzada puede realizarse con de datos de panel.

### 3.1.1.1.- HIPÓTESIS PARA ESTIMAR $\beta$ DE FORMA CONSISTENTE POR EL MÉTODO MCO AGRUPADO.

Siguiendo a Wooldridge (2002: 171), las hipótesis necesarias para que las estimaciones por MCO AGRUPADOS sean consistentes son:

- Hipótesis MCOA.1:  $E(x_t' u_t) = 0 \quad t = 1, 2, \dots, T$

Esta hipótesis, habitual en los modelos MCO, nos dice que no puede haber relación entre  $x_t$  y  $u_t$  en el mismo momento del tiempo, pero sin embargo no nos indica nada acerca de la posible relación que puede producirse entre  $x_s$  y  $u_t$  para  $s \neq t$

- Hipótesis MCOA.2 :  $\text{rango} \left[ \sum_{t=1}^T E(x_t' x_t) \right] = K$

Esta hipótesis es la hipótesis de Rango Pleno o NO MULTICOLINEALIDAD PERFECTA entre las variables explicativas y significa que las variables explicativas no pueden estar “totalmente” correlacionadas entre ellas, ya que esas variables serían las mismas y lógicamente no aportarían nada al modelo.

Para que los estimadores MCO sean eficientes en la regresión MCO agrupada sobre  $i$  y  $t$  necesitamos añadir las hipótesis de homoscedasticidad y no autocorrelación.

- Hipótesis MCOA.3.a



$$E(u_t^2 x_t' x_t) = \sigma^2 E(x_t' x_t) \quad t=1,2,\dots,T \quad \text{donde } \sigma^2 = E[u_t^2] \quad \text{para todo } t$$

- Hipótesis MCOA.3.b

$$E[u_t u_s x_t' x_s] = 0, \quad t \neq s, \quad t, s = 1, 2, \dots, T$$

La hipótesis MCOA.3.a es una hipótesis de homoscedasticidad muy estricta; es suficiente con que  $E[u_t^2 / x_t] = \sigma^2$  para todo  $t$ . En su forma más restrictiva, esta hipótesis, significa no solo que la varianza condicionada no dependa de  $x_t$ , sino que además la varianza incondicionada sea la misma para cada periodo de tiempo.

La hipótesis MCOA.3.b restringe las covarianzas condicionadas de los errores en los diferentes periodos temporales, a cero. El problema que puede plantearse es que cada  $x_t$  siempre contiene al menos una constante, por lo tanto a esta hipótesis MCOA.3.b hay que añadirle que  $E[u_t u_s] = 0$ . Con esto es suficiente para que se cumpla la hipótesis MCOA.3.b:  $E[u_t u_s / x_t x_s] = 0, \quad t \neq s, \quad t, s = 1, 2, \dots, T$  En conclusión la hipótesis MCO.3 quiere decir que las varianzas incondicionadas son constantes y las covarianzas incondicionadas son 0 y que además está restringido a las varianzas y covarianzas condicionadas.

Bajo las hipótesis MCOA.1 y MCOA.3, el estimador de Mínimos Cuadrados Agrupados es:

$$\hat{\beta} = \left( \sum_{i=1}^N \sum_{t=1}^T x_{it}' x_{it} \right)^{-1} \left( \sum_{i=1}^N \sum_{t=1}^T x_{it}' y_{it} \right)$$

Este estimador es consistente y se distribuye asintóticamente como una normal. Si la hipótesis MCOA.3 se mantiene, la varianza asintótica de los estimadores puede estimarse por la expresión:

$$\text{A var}(\hat{\beta}) = \sigma^2 [E(x_t' x_t)]^{-1} / N, \quad \text{por tanto el estimador apropiado para}$$

$$\text{A var}(\hat{\beta}) \text{ es: } \hat{\sigma}^2(x'x)^{-1} = \hat{\sigma}^2 \left( \sum_{i=1}^N \sum_{t=1}^T x'_{it} x_{it} \right)^{-1}$$

Donde  $\hat{\sigma}^2$  es el estimador habitual de la varianza en la regresión agrupada de  $y_{it}$  sobre  $x_{it}$   $t = 1, 2, \dots, T$   $i = 1, 2, \dots, N$

¿Podemos realizar inferencia con estos resultados? De nuevo Wooldridge (2002: 172) nos indica que los test  $t$  y  $F$  habituales que utilizamos en MCO son válidos, el único cambio se produce en el estadístico  $F$  cuando queremos contrastar  $Q$  restricciones lineales en el vector  $\beta$  de dimensión  $K \times 1$  que es:

$$F = \frac{(SSR_r - SSR_{ur})}{SSR_{ur}} \cdot \frac{(NT - K)}{Q}$$

donde  $SSR_{ur}$  es la suma de residuos al cuadrado de la regresión  $y_{it}$  sobre  $x_{it}$  sin imponer ninguna restricción y  $SSR_r$  es la suma de los residuos de la regresión utilizando las  $NT$  observaciones con las restricciones impuestas.

### 3.1.1.2.- MATRIZ DE VARIANZAS ROBUSTA ASINTOTICAMENTE.

Cada vez resulta más habitual en los trabajos empíricos trabajar con estimaciones robustas a la heteroscedasticidad y a la autocorrelación. La ventaja resulta evidente, no es preciso modificar el modelo original para resolver el problema. Esta es la práctica habitual. Si hemos detectado la presencia de estos problemas, lo que hacemos es estimar por algún estimador robusto, de manera que podemos fiarnos de los contrastes y podemos continuar con el trabajo.

Esta práctica es posible realizarla con MCOA, para ello necesitamos un estimador robusto de la varianza del estimador que será válido cuando la hipótesis MCOA.3 no pueda mantenerse.

En este caso la forma de dicha matriz es:

$$A \text{var}(\hat{\beta}) = \left( \sum_{i=1}^N \sum_{t=1}^T x'_{it} x_{it} \right)^{-1} \left( \sum_{i=1}^N \sum_{t=1}^T x'_{it} \hat{u}_{it} \hat{u}'_{it} x_{it} \right) \left( \sum_{i=1}^N \sum_{t=1}^T x'_{it} x_{it} \right)^{-1}$$

Como afirma Wooldridge (2002:175) este estimador es robusto a heteroscedasticidad (tanto condicionada como incondicionada) y arbitrariamente correlacionado en el tiempo (condicionada o incondicionadamente)

### 3.1.1.3.- TEST DE AUTOCORRELACIÓN Y HETEROSCEDASTICIDAD DESPUÉS DE ESTIMAR POR MCO AGRUPADOS.

Trabajar con estimadores robustos, es ponernos la venda antes que la herida, por esta razón son necesarios contraste de autocorrelación y heteroscedasticidad que nos informen de la necesidad de utilizar el estimador robusto. En términos estadísticos, estaríamos trabajando con un estimador que no es eficiente, cuando podemos hacerlo con otro que sí lo es. Para saber cuándo es necesario usar uno u otro necesitamos contrastes que informen de la presencia o ausencia de estos problemas.

- **Test de Autocorrelación**

En el caso de datos de panel, hay una razón más para detectar la presencia de autocorrelación; tal y como recoge Wooldridge (2002: 176). La presencia de autocorrelación supone reconocer que el error de cada periodo contiene un factor omitido constante en el tiempo, que nos lleva a otros métodos de estimación de datos de panel.

Por simplificar trabajamos con un proceso AR(1).

$$u_t = \rho_1 u_{t-1} + e_t \quad \text{donde} \quad E[e_t / x_t, u_{t-1}, x_{t-1}, u_{t-2}, \dots] = 0.$$

Deseamos contrastar la hipótesis nula de que los errores no están serialmente correlacionados frente a la hipótesis alternativa de que el error sigue un proceso AR(1)

La forma de proceder, tal y como Wooldridge (2002:176) propone es, escribir el modelo dinámico bajo la hipótesis de autocorrelación AR(1)

$$y_t = x_t\beta + \rho_1 u_{t-1} + e_t \quad t = 2, 3, \dots, T$$

Si la variable  $u_t$  fuera observable, la manera de proceder sería estimar la ecuación por MCO Agrupados y realizar un test t para  $\rho_1$ . Para implementar este procedimiento en la práctica, reemplazamos  $u_t$  por los residuos MCO Agrupados y realizamos la regresión:

$$y_{it} \text{ sobre } x_{it}, \hat{u}_{i,t-1} \text{ para } t = 2, 3, \dots, T \quad i = 1, 2, \dots, N$$

y hacemos un test t standard sobre el coeficiente de  $\hat{u}_{i,t-1}$ . Este contraste presenta una buena característica, y es que funciona tanto si  $x_i$  es estrictamente exógena como si no.

Una forma distinta del test, válida si asumimos exogeneidad estricta; es calcular el estadístico t en la regresión:

$$\hat{u}_{it} \text{ sobre } \hat{u}_{i,t-1} \quad t = 2, 3, \dots, T \quad i = 1, 2, \dots, N$$

- **Test de Heteroscedasticidad**

La presencia de heteroscedasticidad supone el incumplimiento de la hipótesis MCOA.3a, que es una de las hipótesis necesarias para que los estadísticos habituales que acompañan a la regresión MCO Agrupados sean eficientes.

Asumimos que se cumple  $E[u_t/x_t]=0 \quad t=1,2,\dots,T$  pero, sin embargo no requerimos el cumplimiento de la hipótesis de exogeneidad estricta  $E(x_t'u_t)=0 \quad t=1,2,\dots,T$ .

Siguiendo a Wooldridge (2002: 177) la hipótesis nula de homoscedasticidad puede formularse como  $E[u_t^2/x_t]=\sigma^2 \quad t=1,2,\dots,T$ . Bajo  $H_0$   $u_{it}^2$  está incorrelacionada con cualquier función de  $x_{it}$ ; sea  $h_{it}$  un vector  $1 \times Q$  de funciones no constantes de  $x_{it}$ .

La forma de poner en marcha este contraste es la siguiente:

1. Obtenemos los residuos al cuadrado de la regresión MCO Agrupados:  $\hat{u}_{it}^2$
2. Calculamos el R cuadrado:  $R_c^2$  de la regresión  $\hat{u}_{it}^2$  sobre  $1, h_{it} \quad t=1,2,\dots,t \quad i=1,2,\dots,N$
3. El estadístico es  $NTR_c^2$  que se comporta asintóticamente como una  $\chi^2_Q$  bajo  $H_0$ .

Como nuestro principal interés se centra en el cumplimiento de la hipótesis MCOA.3a se incluyen en  $h_{it}$  los elementos de  $x_{it}$ , sus cuadrados y sus productos cruzados. En otras palabras se trata de una versión del test de White de heteroscedasticidad aplicado en el contexto de datos de panel.

Otra posibilidad, desarrollada por el propio Wooldridge, que tiene como ventaja una pérdida notablemente menor de grados de libertad, es incluir en  $\hat{h}_{it}=(\hat{y}_{it}, \hat{y}_{it}^2)$ , esto es, los valores ajustados de la regresión MCO Agrupados y sus cuadrados.

### 3.1.2.- MODELOS DE DATOS DE PANEL CON EFECTOS NO OBSERVADOS

El modelo lineal de datos de panel que hemos estudiado presentaba la característica de que, como mínimo, el error de cada periodo estaba incorrelacionado con las variables explicativas del mismo periodo.

En ocasiones esta hipótesis resulta demasiado complicada de mantener, especialmente cuando la razón principal de usar datos de panel es la de resolver el

problema de variables omitidas. Variables omitidas, constantes en el tiempo, y que pueden explicar las diferencias entre individuos, empresas, etc.

A continuación nos centraremos en el estudio de modelos que contengan explícitamente estas variables omitidas, también conocidas como efectos NO OBSERVADOS, constantes en el tiempo.

Trataremos a estos efectos no observados, al igual que propone Wooldridge (2002: 247 - 291) como VARIABLES ALEATORIAS. El planteamiento clásico supone tratar a los efectos no observados como parámetros que deben ser estimados. Por lo tanto, y dentro de esta óptica, el tema clave será si los efectos no observados están incorrelacionados o no con las variables explicativas, en lugar de su estimación.

### 3.1.3.- EL PROBLEMA DE LAS VARIABLES OMITIDAS.

Sean  $y$  y  $x \equiv (x_1, x_2, \dots, x_k)$  un conjunto de variables aleatorias observables y sea  $c$  una variable aleatoria no observable, el vector  $(y, x_1, x_2, \dots, x_k, c)$  representa la población de interés.

Imaginemos ahora que estamos interesados en conocer, algo muy habitual en Econometría, los efectos parciales de las variables explicativas observadas  $x_i$  en la función de regresión de la población.

$$E[y/x_1, x_2, \dots, x_k, c]$$

Bajo el supuesto “*ceteris paribus*”, desearíamos mantener constante a  $c$  y obtener los efectos parciales de las variables explicativas observables.

Suponiendo que la relación que une a todas las variables es lineal, tenemos

$$E[Y/X, C] = \beta_0 + X\beta + C$$

nuestro interés recae en el vector de  $K \times 1$  de  $\beta$ , que mide la influencia de las variables  $x_i$  en la variable a explicar. La variable  $c$  es una variable no observable que afecta a  $y$ . Un ejemplo habitual que refleja una situación como la descrita sería explicar el salario en el tiempo de un conjunto de personas en función de variables como pueden ser experiencia en años, años de formación, etc. Es innegable que la habilidad individual afecta al salario, pero esta variable es difícilmente observable. ¿Cómo proceder entonces?

Vamos a comenzar suponiendo que  $c$  no está correlacionada con cada  $x_i$ , esto es, no está relacionada con las variables explicativas cuyo efecto nos interesa. Estadísticamente esto significa decir que  $Cov(x_i, c) \neq 0$  para alguna  $i$

Como la variable  $c$  no es observable cuando se estima el modelo, la variable  $c$  está en el término de error, provocando problemas de estimación muy serios (sesgo de variables omitidas...).

Para resolver este problema en el contexto de datos de panel, vamos a suponer que la variable  $c$  es constante en el tiempo.

La función de regresión poblacional viene dada por

$$E[y_t / x_t c] = \beta_0 + x_t \beta + c \quad t = 1, 2, \dots, T$$

donde  $x_t \beta = \beta_1 x_{t1} + \dots + \beta_k x_{tk}$  y  $x_{ij}$  indica la variación de  $j$  en el momento  $t$ .

Modelos como este presuponen que la variable  $c$  tiene el mismo efecto en la respuesta media de cada periodo de tiempo. Para simplificar y sin perder por ello generalidad, fijamos el coeficiente de  $c$  igual a 1. Tal y como dice Wooldridge (2002:248) “como  $c$  es inobservable, nunca tendremos una unidad natural de medida, no tiene sentido intentar estimar sus efectos parciales”

La hipótesis de que  $c$  es constante en el tiempo (y su efecto parcial también es constante en el tiempo) es vital en el análisis que sigue. Una variable no observada, constante en el tiempo, se llama EFECTO NO OBSERVADO en el análisis de datos de panel. Cuando  $t$  representa diferentes periodos de tiempo para el mismo individuo, el efecto no observado se interpreta como una característica del individuo: habilidad, motivación, etc... Si la unidad de observación son empresas, el efecto no observado contiene características relativas a calidad de dirección, estructura, etc... que pueden contemplarse como constantes en el tiempo.

Antes de pasar a los métodos de estimación de datos de panel, nos detendremos a discutir sobre la naturaleza de los efectos no observados y ciertas características que presentan las variables explicativas observadas.

### 3.1.4.- EFECTOS FIJOS O ALEATORIOS.

El modelo básico de efectos no observados puede escribirse como

$$y_{it} = x_{it}\beta + c_i + u_{it} \quad t = 1, 2, \dots, T \quad i = 1, 2, \dots, N$$

Donde

- $x_{it}$  es una matriz  $1 \times K$  que contiene variables  $K$  observables:
  - que cambian en  $t$  pero no en  $i$ .
  - que cambian en  $i$  pero no en  $t$ .
  - y variables que cambian en  $i$  y en  $t$ .
- $c_i$  es la variable que recoge los efectos no observados; hay muchos otros nombres dados a los efectos no observados: componentes no observados, variables latentes y heterogeneidad no observada.
- $u_{it}$  es llamado error o perturbación idiosincrásica y cambia tanto en  $i$  como en  $t$ .



En la bibliografía existente sobre datos de panel, existe toda una discusión metodológica respecto si la variable no observable  $c_i$  debe ser tratada como efecto aleatorio o como fijo. En un principio el debate se centró sobre si  $c_i$  era una variable aleatoria o un parámetro que debía ser estimado. El nombre actual viene de las aproximaciones tradicionales a los modelos de datos de panel,  $c_i$  recibe el nombre de “efecto aleatorio” cuando se trata como variable aleatoria y “efecto fijo” cuando se trata como parámetro y que debe ser estimado para cada observación  $i$  de sección cruzada.

En la econometría moderna, la discusión está superada pero sin embargo los nombres se mantuvieron, así efectos aleatorios es sinónimo de correlación cero entre las variables explicativas y efectos no observados:  $Cov(x_{it}c_i) = 0 \quad t = 1, 2, \dots, T$ . El término efectos fijos significa que se está permitiendo correlación arbitraria entre el efecto no observado  $c_i$  y las variables explicativas  $x_{it}$ .

### 3.1.5.- ESTIMACIÓN DE MODELOS DE EFECTOS NO OBSERVADOS CON MCO AGRUPADOS.

Bajo ciertas hipótesis, el estimador de MCO Agrupados puede utilizarse para obtener un estimador consistente de  $\beta$  en el modelo

$$y_{it} = x_{it}\beta + c_i + u_{it} \quad t = 1, 2, \dots, T \quad i = 1, 2, \dots, N$$

Escribimos el modelo de la forma siguiente  $y_{it} = x_{it}\beta + v_{it} \quad t = 1, 2, \dots, T$

donde  $v_{it} = c_i + u_{it} \quad t = 1, 2, \dots, T$  es el error compuesto. Para cada  $t$   $v_{it}$  es la suma del efecto NO OBSERVADO y del ERROR IDIOSINCRÁSICO.

La estimación por MCO Agrupados de esta ecuación es consistente si  $E[x'_{it}v_{it}] = 0, \quad t = 1, 2, \dots, T$ . Esto es, que no exista correlación entre las variables explicativas y el error idiosincrásico compuesto ( $v_{it}$ ).

Pero esta NO correlación significa que no hay correlación entre las variables explicativas y el error idiosincrásico ( $u_{it}$ ) y que no exista correlación entre las variables observadas y efecto no observado.

$$E[x'_{it}u_{it}] = 0 \quad y$$

$$E[x'_{it}c_i] = 0 \quad t = 1, 2, \dots, T$$

La hipótesis  $E[x'_{it}v_{it}] = 0$  es una hipótesis restrictiva que obliga a que  $E[x'_{it}u_{it}] = 0$ , hipótesis que se mantiene si hemos modelado con éxito  $E[y_{it}/x_{it}, c_i]$ . Esto es, si el modelo contiene variables estrictamente exógenas es más fácil que esta hipótesis se cumpla, pero si el modelo incluye, como variables explicativas, a variables dependientes retardadas  $y_{i,t-1}$  la hipótesis  $E[x'_{it}v_{it}] = 0$  se viola ya que  $y_{i,t-1}$  y  $c_i$  están correlacionadas.

### 3.1.6.- MÉTODO DE EFECTOS ALEATORIOS.

Al igual que hacíamos con el método MCO Agrupados, en el Método de Efectos Aleatorios se coloca a  $c_i$  en el término de error.

El método de efectos aleatorios impone hipótesis adicionales a las que necesitábamos con el método MCO Agrupados. Estas hipótesis son: exogeneidad estricta y ortogonalidad entre  $c_i$  y  $x_{it}$ .

- Hipótesis EA.1

a)  $E[u_{it}/x_i, c_i] = 0 \quad t = 1, 2, \dots, T \Rightarrow \text{Exogeneidad estricta}$

b)  $E[c_i/x_i] = E[c_i] = 0$  donde  $x_i = (x_{i1}, x_{i2}, \dots, x_{iT}) \Rightarrow \text{Ortogonalidad}$

De estas hipótesis es destacable la hipótesis  $E[c_i / x_i] = E[c_i]$ ;  $E[c_i] = 0$  que nos proporciona un intercepto o constante, que está incluido en  $x_{it}$ , como suele ser habitual en los modelos de regresión lineal.

El método de Efectos Aleatorios utiliza la correlación existente en el error compuesto  $v_{it} = c_i + u_{it}$  mediante el método de Mínimos Cuadrados Generalizados (MCG). Para estar seguros de que estos estimadores de MCG Factibles son consistentes necesitamos que exista exogeneidad estricta entre las variables explicativas y el error compuesto.

Siguiendo nuevamente a Wooldridge (2002: 257), a la hipótesis EA.1 le añadimos:

$$y_{it} = x_{it}\beta + v_{it}$$

$$E[v_{it} / x_i] = 0 \quad t = 1, 2, \dots, T \quad \text{Exogeneidad estricta}$$

donde  $v_{it} = c_i + u_{it}$

Como consecuencia de que el error compuesto y las variables explicativas están incorrelacionados (hipótesis de exogeneidad estricta), es posible aplicar métodos de Mínimos Cuadrados Generalizados que tengan en cuenta la estructura del error:

$$v_{it} = c_i + u_{it}$$

Como tener en cuenta esta estructura del error nos lo muestra Wooldridge (2002: 258 y 259), para ello escribimos el modelo  $y_{it} = x_{it}\beta + v_{it}$  para todos los periodos  $T$  como:

$$y_i = x_i\beta + v_i$$

y  $v_i$  puede escribirse como  $v_i = c_i j_T + u_i$  donde  $j_T$  es un vector de  $T \times 1$  de unos. Con un ejemplo lo veremos más claro. Si tenemos 4 individuos el error compuesto lo escribimos:

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix} = \begin{pmatrix} c_1 & 0 & 0 & 0 \\ 0 & c_2 & 0 & 0 \\ 0 & 0 & c_3 & 0 \\ 0 & 0 & 0 & c_4 \end{pmatrix} \bullet \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}$$

Donde  $c_i$  recoge características intrínsecas de cada individuo.

A continuación definimos la matriz de las varianzas (No condicionadas) de  $v_i$  como  $\Omega = E[v_i v_i']$  que es una matriz de dimensión  $T \times T$  en la que asumimos que es definida positiva. No viene de más recordar que esta matriz es necesariamente la misma para todos los elementos  $i$ .

Para que los estimadores de MCG sean consistentes necesitamos añadir la hipótesis de rango pleno:

- Hipótesis EA.2:  $\text{rango } E[x_i' \Omega^{-1} x_i] = K$

El análisis de efectos aleatorios añade hipótesis adicionales sobre los errores idiosincrásicos que dan a  $\Omega$  una forma especial.

La primera hipótesis es que el error idiosincrásico  $u_{it}$  tiene una varianza no condicionada constante para todo  $t$ .

$$E[u_{it}^2] = \sigma_u^2 \quad t = 1, 2, \dots, T \quad \text{Varianza constante en los individuos y en el tiempo}$$

La segunda hipótesis es que el error idiosincrásico no está autocorrelacionado.

$E[u_{it}u_{is}] = 0 \quad \forall t \neq s$       *Incorrelación entre los individuos en distintos momentos del tiempo*

Con estas dos hipótesis podemos obtener la varianza y covarianza de los elementos de  $v_i$ .

Recordemos que seguimos teniendo presente la hipótesis EA.1.a, por lo que  $E[c_i u_{it}] = 0 \quad t = 1, 2, \dots, T$  y por tanto

$$v_{it} = c_i + u_{it}$$

$$E[v_{it}] = E[c_i + u_{it}] = E[c_i] + E[u_{it}] = 0$$

por la hipótesis EA.1.b

$V[v_{it}] = E[v_{it}^2] = E[(c_i + u_{it})^2] = E[c_i^2] + 2E[c_i u_{it}] + E[u_{it}^2] = \sigma_c^2 + \sigma_u^2$ , que por la hipótesis EA.1.a  $E[c_i u_{it}] = 0$ .

Por lo que  $V[v_{it}] = E[v_{it}^2] = \sigma_c^2 + \sigma_u^2$  y  $\sigma_c^2 = E[c_i^2] \quad \forall t \neq s$

Esta es la varianza, pero queda por definir la covarianza, que viene dada por:

$E[v_{it}v_{is}] = E[(c_i + u_{it})(c_i + u_{is})] = E[c_i^2] + E[c_i u_{is}] + E[c_i u_{it}] + E[u_{it}u_{is}] = E[c_i^2] = \sigma_c^2$  por la hipótesis EA.1.a

Por lo tanto, la matriz  $\Omega$  de varianzas y covarianzas toma la forma de:

$$\Omega = E[v_i v_i'] = \begin{bmatrix} \sigma_c^2 + \sigma_u^2 & \sigma_c^2 & \dots & \sigma_c^2 \\ \sigma_c^2 & \sigma_c^2 + \sigma_u^2 & \dots & \sigma_c^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_c^2 & \dots & \dots & \sigma_c^2 + \sigma_u^2 \end{bmatrix}$$

Si definimos a la matriz  $J_T J_T'$  como una matriz de dimensión  $T \times T$  con unos en cada elemento podemos escribir la matriz de varianzas y covarianzas del error compuesto como

$$\Omega = \sigma_u^2 I_T + \sigma_c^2 J_T J_T'$$

Cuando  $\Omega$  tiene la forma anterior, decimos que  $\Omega$  tiene estructura de efectos aleatorios y sólo depende de dos parámetros  $\sigma_u^2$  y  $\sigma_c^2$  independientemente del tamaño de  $T$ .

Estudiamos ahora la correlación entre  $v_{is}$  y  $v_{it}$ , esto es la correlación entre el mismo individuo en dos momentos distintos de tiempo  $s$  y  $t$ .

$$Corr(v_{is}, v_{it}) = \frac{Cov(v_{is}, v_{it})}{D[v_{is}]D[v_{it}]} = \frac{E[v_{is}v_{it}]}{\sqrt{E[v_{is}^2]}\sqrt{E[v_{it}^2]}} = \frac{\sigma_c^2}{\sqrt{\sigma_c^2 + \sigma_u^2}\sqrt{\sigma_c^2 + \sigma_u^2}} = \frac{\sigma_c^2}{\sigma_c^2 + \sigma_u^2} \geq 0 \quad t \neq s$$

Esta correlación es el cociente entre la varianza de  $c_i$  y la varianza del error compuesto y resulta útil como una medida de la importancia relativa del efecto no observado  $c_i$

Estas hipótesis son especiales para el análisis de Efectos Aleatorios.

Para que los estimadores de MCG Factibles sean eficientes debemos asumir que la matriz de varianzas de  $v_i$  condicionada a  $x_i$  es constante:

$$E[v_i v_i' / x_i] = E[v_i v_i']$$

- Hipótesis EA.3:

$$a) \quad E[u_i u_i' / x_i, c_i] = \sigma_u^2 I_T$$

$$b) \quad E[c_i^2 / x_i] = \sigma_c^2$$

En esta hipótesis están implícitas las hipótesis:  $E[u_{it}^2] = \sigma_u^2$ ,  $E[u_{it} u_{is}] = 0 \quad \forall t \neq s$  y  $E[v_i v_i' / x_i] = E[v_i v_i']$ .

Bajo la hipótesis EA.3 a,  $E[u_{it}^2 / x_i, c_i] = \sigma_u^2$ ,  $t=1,2,\dots,T$  que implica la hipótesis  $E[u_{it}^2] = \sigma_u^2$   $t=1,2,\dots,T$  y  $E[u_{it}u_{is} / x_i, c_i] = 0$   $t \neq s$ ,  $t,s=1,2,\dots,T$  que implica la hipótesis  $E[u_{it}u_{is}] = 0$   $\forall t \neq s$ . Esta hipótesis EA.3 es más fuerte porque asume que las varianzas condicionadas son constantes y las covarianzas condicionadas son 0.

La hipótesis EA.1b:  $E[c_i / x_i] = E[c_i] = 0$  junto con la hipótesis EA.3b es lo mismo que decir que  $Var[c_i / x_i] = Var[c_i]$ , que es la hipótesis de homoscedasticidad sobre los efectos no observados  $c_i$ .

En conclusión, bajo la hipótesis EA.3, la hipótesis  $E[v_i v_i' / x_i] = E[v_i v_i']$  se mantiene y

$$\Omega \text{ tiene la forma } E[v_i v_i'] = \begin{bmatrix} \sigma_c^2 + \sigma_u^2 & \sigma_c^2 & \cdots & \sigma_c^2 \\ \sigma_c^2 & \sigma_c^2 + \sigma_u^2 & \cdots & \sigma_c^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_c^2 & \cdots & \cdots & \sigma_c^2 + \sigma_u^2 \end{bmatrix}.$$

Una vez revisadas las hipótesis sobre las que se asienta el análisis de Efectos Aleatorios, se trata ahora de buscar un procedimiento que nos permita estimar los parámetros del modelo por Mínimos Cuadrados Generalizados Factibles (MCGF). Para implementarlo definimos  $\sigma_v^2 = \sigma_c^2 + \sigma_u^2$  y asumimos que contamos con estimadores consistentes de  $\sigma_c^2$  y  $\sigma_u^2$ .

Estimamos  $\hat{\Omega} = \hat{\sigma}_u^2 I_T + \hat{\sigma}_c^2 J_T J_T'$  como una matriz  $T \times T$  que asumimos como definida positiva.

De forma que el estimador de datos de panel de efectos aleatorios es:

$$\hat{\beta}_{EA} = \left( \sum_{i=1}^N X_i' \hat{\Omega}^{-1} X_i \right)^{-1} \left( \sum_{i=1}^N X_i' \hat{\Omega}^{-1} Y_i \right)$$

Wooldridge (2002: 260) nos dice que este estimador está influido por la hipótesis EA.3 que configura la forma de  $\hat{\Omega}$ . Sin embargo el estimador  $\hat{\beta}_{EA}$  es consistente,

tanto si la hipótesis EA.3 se mantiene como si no. Siempre que se mantenga la hipótesis EA.1 y la condición de rango, el estimador  $\hat{\beta}_{EA} \xrightarrow{P} \beta$  cuando  $N \rightarrow \infty$ .

La hipótesis EA.3 nos sirve para asegurarnos que el estimador de efectos aleatorios es eficiente dentro de la clase de estimadores consistentes.

Como puede observarse, la estimación por MCGF pasa por obtener  $\hat{\Omega}$ , y que a su vez necesita obtener  $\hat{\sigma}_c^2$  y  $\hat{\sigma}_u^2$ , que tal y como afirma Wooldridge (2002: 260) la forma más sencilla será estimar primero  $\hat{\sigma}_v^2 = \hat{\sigma}_c^2 + \hat{\sigma}_u^2$ .

Bajo la hipótesis EA.3.a  $\sigma_v^2 = \sum_{t=1}^T \frac{E[v_{it}^2]}{T}$  para todo  $i$ , y a continuación promediando  $v_{it}^2$  para todo  $i$  y  $t$  obtendríamos un estimador consistente de  $\sigma_v^2$ .

La idea parece sencilla pero ¿cómo obtenemos  $v_{it}^2$ ? En regresión lo habitual es estimar el modelo y trabajar con los residuos, pero para estimar el modelo necesitamos un estimador. ¿Qué estimador utilizar? Wooldridge (2002: 260) dice que un estimador inicial conveniente de  $\beta$  es el estimador MCO Agrupados, que denotaremos por  $\hat{\beta}$  y llamaremos  $\hat{v}_{it}$  a los residuos MCO Agrupados. Un estimador consistente de  $\sigma_v^2$  es:

$\hat{\sigma}_v^2 = \frac{1}{(NT - K)} \sum_{i=1}^N \sum_{t=1}^T \hat{v}_{it}^2$  que es el estimador habitual de la varianza de la regresión de MCO Agrupados.

El paso siguiente es encontrar un estimador consistente de  $\sigma_c^2$ , por las hipótesis de partida, sabemos que  $\sigma_c^2 = E[v_{it}v_{is}] \quad \forall t \neq s$ ; por tanto, para cada  $i$  existen  $\frac{T(T-1)}{2}$  productos de errores no redundantes que pueden utilizarse para estimar  $\sigma_c^2$ . Sumando todas las combinaciones y tomando esperanzas para cada individuo obtenemos:



$$\begin{aligned}
E\left[\sum_{t=1}^{T-1} \sum_{s=t+1}^T v_{it} v_{is}\right] &= \sum_{t=1}^{T-1} \sum_{s=t+1}^T E[v_{it} v_{is}] = \\
&= \sum_{t=1}^{T-1} \sum_{s=t+1}^T \sigma_c^2 = \sigma_c^2 \sum_{t=1}^{T-1} (T-t) = \\
&= \sigma_c^2 [(T-1) + (T-2) + \dots + 2 + 1] = \\
&= \sigma_c^2 \frac{T(T-1)}{2}
\end{aligned}$$

El estimador consistente se obtiene reemplazando las esperanzas por un promedio (en todos los individuos) y reemplazando  $v_{it}$  por los residuos de MCO Agrupados.

$$\hat{\sigma}_c^2 = \frac{1}{\left[ N \frac{T(T-1)}{2} - K \right]} \sum_{i=1}^N \sum_{t=1}^{T-1} \sum_{s=t+1}^T \hat{v}_{it} \hat{v}_{is}$$

Una vez calculados  $\hat{\sigma}_v^2$  y  $\hat{\sigma}_c^2$  podemos calcular  $\hat{\sigma}_u^2 = \hat{\sigma}_v^2 - \hat{\sigma}_c^2$ . Como el estimador de Efectos Aleatorios es un estimador de MCG Factibles, necesitamos estimadores consistentes de  $\sigma_c^2$  y  $\sigma_u^2$  para obtener un estimador  $\sqrt{N}$  eficiente de  $\beta$

Sucede que en la práctica es posible que la estimación  $\hat{\sigma}_c^2$  no sea positiva (en teoría es imposible), aunque en la mayoría de los casos lo es. Un valor negativo de  $\hat{\sigma}_c^2$  nos informa de la existencia de autocorrelación negativa en  $u_{it}$  lo que significa que la hipótesis EA.3.a no se cumple. Si efectivamente  $\hat{\sigma}_c^2$  es negativo, la solución pasa por utilizar MCG Factibles sin restringir.

### 3.1.6.1.- ESTIMADOR ROBUSTO DE LA MATRIZ DE VARIANZAS.

Aunque la hipótesis EA.3 no se cumpla, los estimadores de Efectos Aleatorios continúan siendo consistentes, por lo tanto puede resultar útil algún procedimiento que nos permita realizar inferencia a pesar del incumplimiento de esta hipótesis.

La hipótesis EA.3 puede fallar por dos razones:

1.  $E[v_i v_i' / x_i]$  puede no ser constante, por tanto  $E[v_i v_i' / x_i] \neq E[v_i v_i']$
2.  $E[v_i v_i']$  puede no tener la estructura de efectos aleatorios; el error idiosincrásico  $u_{it}$  puede tener varianzas que cambien en el tiempo o podrían estar correlacionadas.

Wooldridge (2002: 262) nos dice que el procedimiento para realizar inferencia en estos casos es construir una matriz de varianzas robusta. Esta matriz es:

$$\left( \sum_{i=1}^N X_i' \hat{\Omega}^{-1} X_i \right)^{-1} \left( \sum_{i=1}^N X_i' \hat{\Omega}^{-1} \hat{v}_i \hat{v}_i' X_i \right) \left( \sum_{i=1}^N X_i' \hat{\Omega}^{-1} X_i \right)^{-1}$$

donde  $\hat{v}_i = y_i - x_i \hat{\beta}_{EA}$ ,  $i = 1, 2, \dots, N$ .

Los errores estándar robustos (desviaciones típicas robustas) se obtienen de la forma habitual; calculando la raíz cuadrada de los elementos de la diagonal principal de la matriz de varianzas robusta de los estimadores.

El estadístico de Wald robusto para realizar inferencia se obtiene

$$W = (R\hat{\beta} - r)' (R\hat{V}R')^{-1} (R\hat{\beta} - r) \text{ donde } \hat{V} \text{ es la matriz robusta de los estimadores.}$$

Si la hipótesis EA.3 no se cumple, la suma de cuadrados de los residuos para los estadísticos F no es válida y por tanto no tiene sentido realizar cualquier tipo de contraste ya que las conclusiones que vamos a obtener no son válidas. Por esta razón es una buena idea utilizar un análisis robusto cuando sea posible.

### 3.1.6.2.- ANÁLISIS GENERAL DE MÍNIMOS CUADRADOS GENERALIZADOS FACTIBLE (MCGF).

Cuando los errores idiosincrásicos  $\{u_{it} : t=1,2,\dots,T\}$  son heteroscedásticos y están correlacionados en  $t$  Wooldridge (2002:263) nos propone un estimador más general de  $\Omega$ :

$$\hat{\Omega} = \frac{\sum_{i=1}^N \hat{v}_i \hat{v}_i'}{N} \quad \text{donde } \hat{v}_i \text{ son los residuos de MCO Agrupados. Este estimador de MCG}$$

Factibles es consistente bajo las hipótesis EA.1 y EA.2 y si además asumimos que  $E[v_i v_i' / x_i] = \Omega$  resulta entonces que este estimador es asintóticamente eficiente y el estimador de la varianza asintótica tiene la forma habitual.

Este estimador es más general que el utilizado en el análisis de Efectos Aleatorios, entonces ¿por qué no utilizar siempre el estimador de MCG Factibles con

$$\hat{\Omega} = \frac{\sum_{i=1}^N \hat{v}_i \hat{v}_i'}{N} ?$$

El motivo es que la estructura de los datos de

$$\Omega = \begin{bmatrix} \sigma_c^2 + \sigma_u^2 & \sigma_c^2 & \cdots & \sigma_c^2 \\ \sigma_c^2 & \sigma_c^2 + \sigma_u^2 & \cdots & \sigma_c^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_c^2 & \cdots & \cdots & \sigma_c^2 + \sigma_u^2 \end{bmatrix} \text{ es sinónimo de modelos con efectos no}$$

observados; cualquier correlación en el error compuesto  $\{v_{it} : t=1,2,\dots,T\}$  se asume que está causada por la presencia de  $c_i$ . El error idiosincrásico,  $u_{it}$  es por definición homoscedástico y no autocorrelacionado.

### 3.1.7.- MÉTODO DE EFECTOS FIJOS.

Regresemos al modelo lineal de efectos no observados para  $T$  periodos de tiempo.

$$y_{it} = x_{it}\beta + c_i + u_{it} \quad t = 1, 2, \dots, T$$

El método de efectos aleatorios para estimar  $\beta$  colocaba a  $c_i$  en el término de error, bajo la hipótesis de que  $c_i$  era ortogonal a  $x_{it}$  y a continuación consideraba la correlación en el error compuesto  $v_{it} = c_i + u_{it}$  usando un análisis de MCG.

En otras ocasiones, la razón de utilizar datos de panel es precisamente el hecho de permitir que  $c_i$  esté arbitrariamente correlacionado con las  $x_{it}$ .

Los métodos de efectos fijos tienen esta característica.

Las  $T$  ecuaciones del modelo  $y_{it} = x_{it}\beta + c_i + u_{it}$  pueden escribirse como

$$y_i = x_i\beta + c_iJ_T + u_i, \text{ donde } J_T \text{ es un vector de unos de dimensión } T \times 1$$

La primera hipótesis de Efectos Fijos es la de EXOGENEIDAD ESTRICTA de las variables explicativas condicionadas a  $c_i$

- Hipótesis EF.1:  $E[u_{it} / x_i, c_i] = 0 \quad t = 1, 2, \dots, T$

Esta hipótesis es idéntica a la primera parte de la hipótesis EA.1; esto es, mantenemos la exogeneidad estricta de  $\{x_{it} : t = 1, 2, \dots, T\}$  condicionada a los efectos no observados. La diferencia fundamental radica en que no necesitamos la hipótesis EA.1b. Dicho de otra forma, en el análisis de efectos fijos la relación entre  $c_i$  y  $x_i$  está permitida para cualquier valor de  $x_i$ . Esto es,  $E[c_i / x_i] \neq 0$

Tal y como nos dice Wooldridge (2002: 266), al relajar la hipótesis EA.1b podemos estimar de forma consistente los efectos parciales en presencia de variables omitidas constantes en el tiempo que puedan estar arbitrariamente relacionadas con las variables observadas  $x_{it}$ . En consecuencia, el análisis de Efectos Fijos es más robusto que el análisis de Efectos Aleatorios; pero esta robustez tiene un precio. No podemos incluir variables constantes en el tiempo en  $x_{it}$ . La razón es simple, si  $c_i$  puede estar arbitrariamente correlacionada con cada elemento de  $x_{it}$ , no hay forma de distinguir los efectos de las variables observables constantes en el tiempo de los efectos de la variable no observable constante en el tiempo  $c_i$ .

La manera de estimar  $\beta$  bajo la hipótesis EF.1 es transformar la ecuación de forma que eliminemos los efectos no observados  $c_i$ . Ahora veremos la TRANSFORMACIÓN DE EFECTOS FIJOS, llamada también transformación WITHIN.

La transformación EF se obtiene promediando la ecuación  $y_{it} = x_{it}\beta + c_i + u_{it}$  en el periodo  $t = 1, 2, \dots, T$  para obtener la ecuación de sección cruzada.

$$\bar{y}_i = \bar{x}_i\beta + c_i + \bar{u}_i$$

donde  $\bar{y}_i = \frac{\sum_{t=1}^T y_{it}}{T}$ ,  $\bar{x}_i = \frac{\sum_{t=1}^T x_{it}}{T}$  y  $\bar{u}_i = \frac{\sum_{t=1}^T u_{it}}{T}$ ; a continuación restamos

$\bar{y}_i = \bar{x}_i\beta + c_i + \bar{u}_i$  de la ecuación  $y_{it} = x_{it}\beta + c_i + u_{it}$  y para cada  $t$  obtenemos la ecuación EF transformada.

$$y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)\beta + u_{it} - \bar{u}_i$$

o

$$\ddot{y}_{it} = \ddot{x}_{it}\beta + \ddot{u}_{it} \quad t = 1, 2, \dots, T$$

donde  $\ddot{y}_{it} = y_{it} - \bar{y}_i$ ,  $\ddot{x}_{it} = x_{it} - \bar{x}_i$  y  $\ddot{u}_{it} = u_{it} - \bar{u}_i$

La media temporal de la ecuación original ha eliminado el efecto individual específico  $c_i$ .

Con  $c_i$  fuera de la ecuación, lo lógico es pensar en estimar la ecuación  $\ddot{y}_{it} = \ddot{x}_{it}\beta + \ddot{u}_{it}$  por MCO Agrupados.

Para saber si esta estimación es consistente, necesitamos demostrar que la hipótesis clave de los MCO Agrupados se mantiene en la ecuación  $\ddot{y}_{it} = \ddot{x}_{it}\beta + \ddot{u}_{it}$ . Esto es,

$$E[\ddot{x}'_{it}\ddot{u}_{it}] = 0 \quad t = 1, 2, \dots, T$$

Para cada  $t$ , el lado izquierdo de la ecuación puede escribirse como  $E[(x_{it} - \bar{x}_i)'(u_{it} - \bar{u}_i)]$ . Ahora, bajo la hipótesis EF.1,  $u_{it}$  está incorrelacionada con  $x_{is}$ , para todo  $s, t = 1, 2, \dots, T$ . Por lo tanto,  $u_{it}$  y  $\bar{u}_i$  están incorrelacionadas con  $x_{it}$  y  $\bar{x}_i$  para  $t = 1, 2, \dots, T$ . Por lo que la hipótesis  $E[\ddot{x}'_{it}\ddot{u}_{it}] = 0$  se mantiene bajo la hipótesis EF.1 y por tanto, la utilización de MCO Agrupados en la ecuación  $\ddot{y}_{it} = \ddot{x}_{it}\beta + \ddot{u}_{it}$  producirá estimadores consistentes.

Si analizamos con detalle, tal como hace Wooldridge (2002: 268) la expresión  $E[\ddot{x}'_{it}\ddot{u}_{it}] = 0$ , bajo la hipótesis EF.1 se cumple que

$$E[\ddot{u}_{it} / x_i] = E[(u_{it} - \bar{u}_i) / x_i] = E[u_{it} / x_i] - E[\bar{u}_i / x_i] = 0 \quad \text{que implica que}$$

$$E[\ddot{u}_{it} / \ddot{x}_{i1}, \ddot{x}_{i2}, \dots, \ddot{x}_{iT}] = 0, \text{ donde cada } \ddot{x}_{it} \text{ es una función de } x_i = (x_{i1}, x_{i2}, \dots, x_{iT}).$$

Estos resultados demuestran que  $\ddot{x}_{it}$  satisface la forma de esperanza condicionada de la hipótesis de exogeneidad estricta en el modelo  $\ddot{y}_{it} = \ddot{x}_{it}\beta + \ddot{u}_{it}$ . Esta conclusión implica que el estimador de efectos fijos de  $\beta$  es insesgado bajo la hipótesis EF.1.

Por lo tanto el estimador de EFECTOS FIJOS (EF), denotado por  $\hat{\beta}_{EF}$ , es el estimador de MCO Agrupados de la regresión

$\ddot{y}_{it}$  sobre  $\ddot{x}_{it}$   $t = 1, 2, \dots, T, i = 1, 2, \dots, N$

El estimador EF es sencillo de calcular una vez que hemos obtenido la media en el tiempo.

Para expresarlo matricialmente matricial procedemos de la siguiente manera. Escribimos la expresión  $\ddot{y}_{it} = \ddot{x}_{it}\beta + \ddot{u}_{it}$  para todos los periodos de tiempo como

$$\ddot{y}_i = \ddot{x}_i\beta + \ddot{u}_i$$

donde  $\ddot{y}_i$  es un vector de dimensión  $T \times 1$ ,  $\ddot{x}_i$  es una matriz de dimensión  $T \times k$  y  $\ddot{u}_i$  un vector de dimensión  $T \times 1$ . Este conjunto de ecuaciones se obtiene pre multiplicando la ecuación  $y_i = x_i\beta + c_iJ_T + u_i$  por una matriz de media temporal; para ello definimos  $Q_T = I_T - J_T(J_T'J_T)^{-1}J_T'$  que es una matriz  $T \times T$  simétrica, idempotente y con rango  $T-1$ . Además

$$Q_TJ_T = 0, Q_TY_i = \ddot{Y}_i, Q_TX_i = \ddot{X}_i \text{ y } Q_TU_i = \ddot{U}_i,$$

cuando pre multiplicamos la ecuación  $y_i = x_i\beta + c_iJ_T + u_i$  por  $Q_T$  obtenemos la ecuación  $\ddot{y}_i = \ddot{x}_i\beta + \ddot{u}_i$

Para asegurarnos que el estimador de Efectos Fijos tiene un buen comportamiento asintótico, imponemos la condición habitual de rango sobre la matriz de variables explicativas una vez restada la media temporal

- Hipótesis EF.2: 
$$rango\left[\sum_{t=1}^T E[\ddot{x}_{it}'\ddot{x}_{it}]\right] = rango\left[E[\ddot{X}_i'\ddot{X}_i]\right] = K$$

Si  $x_{it}$  contiene un elemento que no varía en el tiempo para cada  $i$ , el correspondiente elemento en  $\ddot{x}_{it}$  es 0 para todo  $t$  y para cualquier sección cruzada. Como  $\ddot{X}_i$  podría

contener una columna de ceros para alguna  $i$ , entonces la hipótesis EF.2 podría no ser cierta. En otras palabras, la hipótesis EF.2 demuestra de una forma explícita por qué las variables constantes en el tiempo no están permitidas en el análisis de efectos fijos.

El estimador de efectos fijos puede expresarse como

$$\hat{\beta}_{FE} = \left( \sum_{i=1}^N \ddot{X}_i' \ddot{X}_i \right)^{-1} \left( \sum_{i=1}^N \ddot{X}_i' \ddot{y}_i \right) = \left( \sum_{i=1}^N \sum_{t=1}^T \ddot{x}_{it}' \ddot{x}_{it} \right)^{-1} \left( \sum_{i=1}^N \sum_{t=1}^T \ddot{x}_{it}' \ddot{y}_{it} \right)$$

Este estimador es también conocido como estimador WITHIN porque utiliza las variaciones temporales dentro de cada sección cruzada.

Otro estimador famoso es el estimador BETWEEN. Este último utiliza exclusivamente la variación entre las observaciones de sección cruzada, y es el resultante de aplicar el estimador MCO a la ecuación

$$\bar{y}_i = \bar{x}_i \beta + c_i + \bar{u}_i$$

$$\text{donde } \bar{y}_i = \frac{\sum_{t=1}^T y_{it}}{T} \quad \bar{x}_i = \frac{\sum_{t=1}^T x_{it}}{T} \quad \bar{u}_i = \frac{\sum_{t=1}^T u_{it}}{T}$$

Este estimador no es consistente bajo la hipótesis EF.1 porque  $E[\bar{x}_i' c_i]$  no es necesariamente 0. El estimador BETWEEN es consistente bajo la hipótesis EA.1 y condición de rango standard, pero descarta la información de las series de datos temporales en el conjunto de datos. Resulta más eficiente utilizar el estimador de efectos aleatorios.

Para poder realizar inferencia con el estimador de Efectos Fijos necesitamos añadir una hipótesis adicional. Si esta hipótesis se cumple el estimador EF será un estimador eficiente.



- Hipótesis EF.3:  $E[u_i u_i' / x_i, c_i] = \sigma_u^2 I_T$

Esta hipótesis es idéntica a la hipótesis EA.3a.

Gracias a la hipótesis EF.1 podemos decir que  $E[u_i / x_i, c_i] = 0$  por lo que la hipótesis EF.3 es lo mismo que decir  $Var[u_i / x_i, c_i] = \sigma_u^2 I_T$ , siempre y cuando la hipótesis EF.1 se mantenga.

Como hicimos en otro momento, vamos a descomponer la hipótesis EF.3 en dos partes.

- La primera es  $E[u_i u_i' / x_i, c_i] = E[u_i u_i']$  que es bastante habitual en el contexto de métodos de estimación.
- La segunda es que la matriz de varianzas No condicionada  $E[u_i u_i']$  tiene la forma especial  $\sigma_u^2 I_T$ . Esto implica que el error idiosincrásico  $u_{it}$  tiene la varianza constante en  $t$  y no está autocorrelacionado.

La hipótesis EF.3 junto con la hipótesis EF.1 nos llevan a que la matriz de varianzas No condicionadas del error compuesto  $v_i = c_i J_T + u_i$  tiene la misma forma que la varianza de los efectos aleatorios.

No resulta tan evidente que la hipótesis EF.3 tenga como consecuencia la de asegurarnos la eficiencia de los estimadores de efectos fijos y guiarnos hacia un cálculo sencillo de los errores estándar y de los test estadísticos. Wooldridge (2002: 270) nos muestra el camino. Para ello vamos a trabajar con la ecuación una vez restada la media:

$$\ddot{y}_{it} = \ddot{x}_{it} \beta + \ddot{u}_{it} \quad t = 1, 2, \dots, T$$

para que los estimadores de MCO Agrupados sean relativamente eficientes, necesitamos que la matriz de varianzas y covarianzas del error idiosincrásico  $\{\ddot{u}_{it} : t=1,2,\dots,T\}$  sean homoscedástica en  $t$  y no presente autocorrelación.

La varianza de  $\ddot{u}_{it}$  puede calcularse como

$$E[\ddot{u}_{it}^2] = E[(u_{it} - \bar{u}_i)^2] = E[u_{it}^2] + E[\bar{u}_i^2] - 2E[u_{it}\bar{u}_i] = \sigma_u^2 + \frac{\sigma_u^2}{T} - 2\frac{\sigma_u^2}{T} = \sigma_u^2 \left(1 - \frac{1}{T}\right)$$

ya que

$$E[\bar{u}_i^2] = E\left[\left(\frac{\sum_{t=1}^T u_{it}}{T}\right)^2\right] = \frac{1}{T^2} \sum_{t=1}^T E[u_{it}^2] = \frac{T}{T^2} \sigma_u^2 = \frac{\sigma_u^2}{T}$$

$$E[u_{it}\bar{u}_i] = E\left[u_{it} \frac{\sum_{t=1}^T u_{it}}{T}\right] = \frac{1}{T} E[u_{it}(u_{i1} + u_{i2} + \dots + u_{iT})] =$$

$$\frac{1}{T} [E[u_{it}u_{i1}] + E[u_{it}u_{i2}] + \dots + E[u_{it}u_{iT}]] = \frac{1}{T} [0 + 0 + \dots + E[u_{it}^2]] = \frac{\sigma_u^2}{T}$$

que verifica de forma No condicionada la homoscedasticidad en  $t$ . Sin embargo, para  $t \neq s$  la covarianza entre  $\ddot{u}_{it}$  y  $\ddot{u}_{is}$  es

$$E[\ddot{u}_{it}\ddot{u}_{is}] = E[(u_{it} - \bar{u}_i)(u_{is} - \bar{u}_i)] = E[u_{it}u_{is}] - E[u_{it}\bar{u}_i] - E[u_{is}\bar{u}_i] + E[\bar{u}_i^2] =$$

$$0 - \frac{\sigma_u^2}{T} - \frac{\sigma_u^2}{T} + \frac{\sigma_u^2}{T} = -\frac{\sigma_u^2}{T} < 0$$

Calculamos el coeficiente de correlación a partir de lo anterior:

$$Corr(\ddot{u}_{it}, \ddot{u}_{is}) = \frac{Cov(\ddot{u}_{it}, \ddot{u}_{is})}{\sqrt{aVar(\ddot{u}_{it})}\sqrt{aVar(\ddot{u}_{is})}} = \frac{E[\ddot{u}_{it}\ddot{u}_{is}]}{\sqrt{E[\ddot{u}_{it}^2]}\sqrt{E[\ddot{u}_{is}^2]}} = \frac{\frac{-\sigma_u^2}{T}}{\sigma_u^2(1-\frac{1}{T})} = -\frac{1}{(T-1)}$$

que demuestra que los errores  $\ddot{u}_{it}$  están autocorrelacionados de forma negativa (cuando  $T$  es grande, la autocorrelación tiende a 0).

A continuación necesitamos encontrar la varianza asintótica de los estimadores de efectos fijos  $\hat{\beta}_{EF}$  para poder realizar la inferencia que precisamos; para ello escribimos:

$$\sqrt{N}(\hat{\beta}_{FE} - \beta) = \left( \frac{\sum_{i=1}^N \ddot{X}_i' \ddot{X}_i}{N} \right)^{-1} \left( \frac{\sum_{i=1}^N \ddot{X}_i' u_i}{N} \right) \quad \text{donde hemos utilizado el hecho de que}$$

$$\ddot{X}_i' \ddot{u}_i = X_i' Q_T u_i = \ddot{X}_i' u_i.$$

Bajo la hipótesis EF.3;  $E[u_i u_i' / \ddot{X}_i] = \sigma_u^2 I_T$ ; de manera que aplicando lo que ya

conocemos  $\sqrt{N}(\hat{\beta}_{EF} - \beta) \square Normal\left(0, \sigma_u^2 E[\ddot{X}_i' \ddot{X}_i]^{-1}\right)$  y por lo tanto

$$A \text{ var}(\hat{\beta}_{EF}) = \frac{\sigma_u^2 [E[\ddot{X}_i' \ddot{X}_i]]^{-1}}{N}$$

Sea  $\hat{\sigma}_u^2$  un estimador consistente de  $\sigma_u^2$ , la ecuación anterior es posible estimarla

reemplazando  $E[\ddot{X}_i' \ddot{X}_i]$  por su media muestral  $\frac{\sum_{i=1}^N \ddot{X}_i' \ddot{X}_i}{N}$ , de tal forma que la expresión definitiva que podemos estimar es:

$$A \text{ var}(\hat{\beta}_{EF}) = \hat{\sigma}_u^2 \left( \sum_{i=1}^N \ddot{X}_i' \ddot{X}_i \right)^{-1} = \hat{\sigma}_u^2 \left( \sum_{i=1}^N \sum_{t=1}^T \ddot{x}_{it}' \ddot{x}_{it} \right)^{-1}$$

Los errores estándar asintóticos de las estimaciones de efectos fijos se obtienen de las raíces cuadradas de los elementos de la diagonal principal de la matriz.

La expresión  $\text{var}(\hat{\beta}_{EF}) = \hat{\sigma}_u^2 \left( \sum_{i=1}^N \ddot{X}_i' \ddot{X}_i \right)^{-1} = \hat{\sigma}_u^2 \left( \sum_{i=1}^N \sum_{t=1}^T \ddot{x}_{it}' \ddot{x}_{it} \right)^{-1}$  es muy adecuada ya que presenta una gran similitud con la matriz de varianzas del estimador de MCO que pudiera obtenerse de la regresión de MCO Agrupados de  $\ddot{y}_{it} = \ddot{x}_{it} \beta + \ddot{u}_{it} \quad t=1,2,\dots,T$ .

La dificultad está en la obtención del estimador  $\hat{\sigma}_u^2$  de  $\sigma_u^2$ . Para estimar  $\sigma_u^2$  usamos la ecuación  $E[\ddot{u}_{it}^2]$  sumando a lo largo de  $t$ ;  $\sum_{t=1}^T E[\ddot{u}_{it}^2] = (T-1)\sigma_u^2$  y por tanto

$$\sigma_u^2 = \frac{\sum_{i=1}^N \sum_{t=1}^T E[\ddot{u}_{it}^2]}{N(T-1)}.$$

A continuación definimos los residuos de la estimación de efectos fijos como

$$\hat{u}_{it} = \ddot{y}_{it} - \ddot{x}_{it} \hat{\beta}_{EF} \quad t=1,2,\dots,T \quad i=1,2,\dots,N$$

que son los residuos MCO de la regresión agrupada  $\ddot{y}_{it} = \ddot{x}_{it} \beta + \ddot{u}_{it} \quad t=1,2,\dots,T$ . Entonces el estimador consistente de  $\sigma_u^2$  bajo las hipótesis EF.1 a EF.3 es

$$\hat{\sigma}_u^2 = \frac{\sum_{i=1}^N \sum_{t=1}^T \hat{u}_{it}^2}{N(T-1) - K} = \frac{SSR}{N(T-1) - K}$$

En algunas ocasiones puede ser interesante realizar una estimación de  $\sigma_c^2$ ; esta estimación es útil para determinar la cuantía de la varianza de los componentes no observados respecto a la varianza del componente idiosincrásico.

$$\text{Dado } \hat{\beta}_{EF}, \quad \hat{\sigma}_v^2 = \frac{\sum_{i=1}^N \sum_{t=1}^T (y_{it} - x_{it} \hat{\beta}_{EF})^2}{(NT - K)}$$

es un estimador consistente de  $\sigma_v^2 = \sigma_c^2 + \sigma_u^2$  y por lo tanto un estimador consistente de  $\sigma_c^2$  es  $\hat{\sigma}_v^2 - \hat{\sigma}_u^2$ .

Si deseamos contrastar restricciones múltiples lo más sencillo es utilizar un estadístico F con los grados de libertad adecuadamente calculados. Sea  $SSR_{ur}$  la suma de residuos al cuadrado no restringida de la regresión  $\ddot{y}_{it}$  sobre  $\ddot{x}_{it}$  y sea  $SSR_u$  la suma de residuos al cuadrado restringida de una regresión similar, pero con  $Q$  restricciones impuestas sobre  $\beta$ . Entonces el estadístico de contraste toma la expresión siguiente:

$$F = \frac{(SSR_r - SSR_{ur})}{SSR_{ur}} \cdot \frac{N(T-1) - K}{Q}$$
 que se distribuye como una  $F$  con  $Q$  y  $N(T-1)-K$  grados de libertad,

### 3.1.7.1.- REGRESIÓN DE LA VARIABLE FICTICIA.

Hemos considerado a  $c_i$  como una variable aleatoria no observable, y en la mayor parte de las aplicaciones, esta aproximación proporciona la interpretación apropiada de  $\beta$ . Las aproximaciones tradicionales a la estimación de efectos fijos contemplan a  $c_i$  como un parámetro a ser estimado junto con  $\beta$ .

Nos interesa ahora, acercarnos a este planteamiento y estimar esta variable no observable. Para ello cambiamos la hipótesis EF.2 a su versión para muestras finitas,  $\text{rango}(\ddot{X}\ddot{X}) = K$  entonces el modelo bajo las hipótesis EF.1 a EF.3 satisface las hipótesis de Gauss-Markov condicionadas a  $X$ .

Si  $c_i$  es un parámetro a estimar ¿cómo estimaríamos cada  $c_i$  junto con  $\beta$ ?

Una posibilidad es definir  $N$  variables ficticias, una para cada observación de sección cruzada y correr la regresión de MCO Agrupados.

$$y_{it} \text{ sobre } d_{1i}, d_{2i}, \dots, d_{Ni}, x_{it} \quad t = 1, 2, \dots, T \quad i = 1, 2, \dots, N$$

Llamamos  $\hat{C}_1$  al coeficiente de  $d_{1i}$ ,  $\hat{C}_2$  al coeficiente de  $d_{2i}$  y así sucesivamente. El estimador de  $\beta$  obtenido de esta regresión es el estimador de efectos fijos y esta es la razón por la que  $\hat{\beta}_{EF}$  es llamado el estimador de variable ficticia. Además, los residuos de esta regresión son idénticos a los residuos de la regresión de  $\ddot{y}_{it}$  sobre  $\ddot{x}_{it}$ .

Una de las ventajas de la regresión  $y_{it}$  sobre  $d_{1i}, d_{2i}, \dots, d_{Ni}, x_{it}$  es que produce la estimación apropiada de  $\sigma_u^2$  porque usa  $NT - N - K = N(T-1) - K$  como grados de libertad.

La diferencia más importante entre  $\hat{C}_i$  y  $\hat{\beta}_{EF}$  es que  $\hat{\beta}_{EF}$  es consistente con  $T$  fijo y  $N \rightarrow \infty$ . Este no es el caso con  $\hat{C}_i$ . Cada vez que una nueva sección cruzada se añade, otro  $C_i$  se añade y la información no se acumula en  $C_i$  cuando  $N \rightarrow \infty$ . Cada  $\hat{C}_i$  es un estimador insesgado de  $C_i$  cuando  $C_i$  es tratado como parámetro, pero para ello debemos mantener la hipótesis EF.1 y la de muestras finitas análoga a la hipótesis EF.2. Si añadimos la hipótesis EF.3, manteniendo la hipótesis de Gauss – Markov (condicionada a  $X$ )  $\hat{C}_1, \hat{C}_2, \dots, \hat{C}_N$  son los mejores estimadores lineales insesgados condicionados a  $X$  (es un ejemplo de estimador insesgado pero no consistente).

El software econométrico que emplea efectos fijos suprime habitualmente las estimaciones de  $C_i$ , sin embargo suele ofrecerse un intercepto global. Este intercepto global es una unidad arbitraria de sección cruzada o, más comúnmente, el promedio de  $\hat{C}_i$  para  $i$ .

En ocasiones resulta útil obtener  $\hat{C}_i$  incluso cuando la regresión  $y_{it}$  sobre  $d_{1i}, d_{2i}, \dots, d_{Ni}, x_{it}$   $t=1, 2, \dots, T$   $i=1, 2, \dots, N$  no es factible. En estos casos, utilizando la primera condición de MCO, cada  $\hat{C}_i$  es estimado por  $\hat{C}_i = \bar{Y}_i - \bar{X}_i \hat{\beta}_{EF}$   $i=1, 2, \dots, N$

Una vez calculados los  $\hat{C}_i$ , podemos calcular la media muestral, la desviación típica muestral y los cuantiles para tener una idea de la heterogeneidad existente en la población.

Bajo las hipótesis del modelo lineal clásico (se requieren las hipótesis EF.1 – EF.3 y además la hipótesis de normalidad) podemos realizar un contraste sobre la igualdad de  $C_i$  utilizando un test  $F$  estándar para  $T$  o cualquier tamaño (los grados de libertad son  $N-1$  y  $N(T-1)-K$ ). Desgraciadamente las propiedades de este test con  $N \rightarrow \infty$  y con  $T$  fijo son desconocidas sin la hipótesis de normalidad.

### 3.1.7.2.- AUTOCORRELACIÓN Y ESTIMADOR ROBUSTO DE LA MATRIZ DE VARIANZAS.

El estimador de Efectos Fijos es consistente y asintóticamente normal bajo las hipótesis EF.1 y EF.2. Ahora bien, si la hipótesis EF.3 no se cumple, la expresión

propuesta para estimar la varianza:  $\hat{\sigma}_u^2 \left( \sum_{i=1}^N \sum_{t=1}^T \ddot{x}_{it}' \ddot{x}_{it} \right)^{-1}$  proporciona un estimador

inadecuado de la matriz de varianzas.

En estos casos, la heteroscedasticidad en la perturbación aleatoria representa siempre un problema potencial, pero en ocasiones la autocorrelación pudiera ser un problema mayor. Tal y como nos apunta Wooldridge (2002:274) cuando calculamos el estimador de Efectos Fijos no estamos eliminando la autocorrelación que pudiera existir en  $\{u_{it} \quad t=1,2,\dots,T\}$ ; además, aunque la autocorrelación observada en el error compuesto:  $v_{it} = c_i + u_{it}$  suele estar dominada por la presencia de  $c_i$ , también pudiera existir autocorrelación que se mantuviera en el tiempo.

Contrastar la presencia de autocorrelación en el error idiosincrásico  $\{u_{it}\}$  es complicado; ya que no podemos estimar  $u_{it}$ ; porque la media temporal utilizada en la

estimación de Efectos Fijos nos permite estimar exclusivamente  $\ddot{u}_{it}$ . La única pista que tenemos es que los errores  $\ddot{u}_{it}$  están correlacionados negativamente si los  $u_{it}$  están incorrelacionados  $Corr(\ddot{u}_{it}, \ddot{u}_{is}) = -\frac{1}{(T-1)}$ .

La formulación del contraste resulta complicada ya que el error  $\ddot{u}_{it}$  presenta autocorrelación bajo la hipótesis nula.

Hipótesis nula:  $H_o: \delta = \frac{-1}{T-1}$  donde  $\delta$  es la correlación  $(\ddot{u}_{i,t-1}, \ddot{u}_{i,t})$  y que tomará dicho valor exclusivamente cuando la correlación entre  $(u_{i,t-1}, u_{i,t})$  sea 0; esto es, cuando el modelo no presente autocorrelación.

Para realizar este contraste tenemos dos posibilidades:

La primera, es usar sólo dos periodos de tiempo (por ejemplo, los dos últimos) para contrastar la expresión  $Corr(\ddot{u}_{it}, \ddot{u}_{is}) = -\frac{1}{(T-1)}$ . La forma de proceder es realizar la regresión simple de  $\hat{u}_{i,T}$  sobre  $\hat{u}_{i,T-1}$   $i=1, 2, \dots, N$  y usamos como  $\hat{\delta}$  el coeficiente de  $\hat{u}_{i,T-1}$  y su error estándar para contrastar  $H_o: \delta = \frac{-1}{T-1}$  donde  $\delta$  es el coeficiente de correlación entre  $\hat{u}_{i,T}$   $\hat{u}_{i,T-1}$ . Bajo las hipótesis EF.1 a EF.3 el estadístico  $t$  habitual sigue una distribución asintóticamente normal.

La segunda posibilidad es utilizar más periodos de tiempo si queremos que el estadístico  $t$  sea robusto a autocorrelación arbitraria. Para ello, hacemos la regresión por Mínimos Cuadrados Ordinarios de  $\hat{u}_{i,t}$  sobre  $\hat{u}_{i,t-1}$   $t=1, 2, \dots, T$   $i=1, 2, \dots, N$  y a continuación utilizamos los errores estándar robustos para los MCO Agrupados.

Parece un tanto absurdo que a un test para detectar la autocorrelación le hagamos robusto a correlación. El problema radica en la hipótesis nula: los errores  $\ddot{u}_{i,t}$  están autocorrelacionados y si lo están, esto implica que  $u_{it}, u_{is}$  están incorrelacionados.



Si rechazamos la hipótesis nula, deberíamos como mínimo ajustar el estimador asintótico de la matriz de varianzas y los contrastes estadísticos para conseguir un estimador robusto de la matriz de varianzas.

La forma de ajustar este estimador asintótico de la varianza es obtener  $\hat{u}_i = \ddot{y}_i - \ddot{x}_i \hat{\beta}_{EF}$   $i = 1, 2, \dots, N$  que es un vector  $T \times 1$  de residuos de efectos fijos. El estimador robusto de la matriz de varianzas de los estimadores de efectos fijos es:

$$Av\hat{ar}(\hat{\beta}_{EF}) = (\ddot{X}\ddot{X})^{-1} \left( \sum_{i=1}^N \ddot{x}_i' \hat{u}_i \hat{u}_i' \ddot{x}_i \right) (\ddot{X}\ddot{X})^{-1}$$

Este estimador robusto fue sugerido por Arellano (1987) siguiendo los resultados generales de White (1984 o en la edición revisada del año 2000).

Este estimador robusto de la matriz de varianzas es válido en presencia de cualquier tipo de heteroscedasticidad o autocorrelación en  $\{u_{it} : t = 1, 2, \dots, T\}$  cuando  $T$  es relativamente pequeña respecto a  $N$ . Los errores estándar robustos se obtienen calculando las raíces cuadradas de los elementos de la diagonal de la matriz

$$Av\hat{ar}(\hat{\beta}_{FE}) = (\ddot{X}\ddot{X})^{-1} \left( \sum_{i=1}^N \ddot{x}_i' \hat{u}_i \hat{u}_i' \ddot{x}_i \right) (\ddot{X}\ddot{X})^{-1}.$$

Además esta matriz puede usarse como la matriz  $\hat{V}$  en la construcción del estadístico de Wald.

El problema se plantea cuando la hipótesis EF.3 no se cumple, en ese caso la forma de la suma de residuos al cuadrado del estadístico  $F$  no es válida asintóticamente.

### 3.1.7.3.- EFECTOS FIJOS CON MÍNIMOS CUADRADOS GENERALIZADOS (MCG).

Wooldridge (2002: 276) nos dice que la hipótesis EF. 3 puede fallar por dos razones.

La primera es que la matriz de la varianza condicionada no sea igual a la matriz de varianza no condicionada:  $E[u_i u_i' / x_i, c_i] \neq E[u_i u_i']$

La segunda es que incluso si  $E[u_i u_i' / x_i, c_i] = E[u_i u_i']$ , la matriz de varianzas no condicionadas puede no ser un escalar:  $E[u_i u_i'] \neq \sigma_u^2 I_T$ , lo que significa que la varianza de  $u_{it}$  cambia en  $t$  o, probablemente más importante, que hay autocorrelación en el error idiosincrásico. En ambos casos la matriz de varianza robusta  $Avâr(\hat{\beta}_{FE}) = (\ddot{X}'\ddot{X})^{-1} \left( \sum_{i=1}^N \ddot{x}_i' \hat{u}_i \hat{u}_i' \ddot{x}_i \right) (\ddot{X}'\ddot{X})^{-1}$  sigue siendo válida.

Una solución distinta a la de calcular la matriz de varianzas robusta para el estimador de Efectos fijos es relajar la hipótesis EF.3 para permitir una matriz de covarianzas condicionadas no restringida y al menos constante. Este es un camino a seguir cuando los errores estándar robusto de los estimadores de efectos fijos son demasiado grandes para ser útiles y además hay evidencias de autocorrelación o de que la varianza de  $u_{it}$  varía en el tiempo.

Hipótesis EFMCG.3:  $E[u_i u_i' / x_i, c_i] = \wedge$  una matriz  $T \times T$  definida positiva.

Bajo la hipótesis EFMCG.3,  $E[\ddot{u}_i \ddot{u}_i' / \ddot{x}_i] = E[\ddot{u}_i \ddot{u}_i']$  y como  $\ddot{u}_i = Q_T u_i$  nos queda entonces la expresión:  $E[\ddot{u}_i \ddot{u}_i'] = Q_T E[u_i u_i'] Q_T = Q_T \wedge Q_T$

que tiene rango  $T-1$ . El rango de esta última matriz provoca problemas en la aproximación habitual de MCG, porque la matriz de varianzas no puede invertirse. Un camino es utilizar la inversa generalizada, pero otra vía más sencilla (y

algebraicamente idéntica) es eliminar uno de los periodos de tiempo del análisis. No importa cuál sea el período a eliminar, el resultado del estimador de MCG es el mismo, tal y como lo demuestran Im, Ahn, Schmidt y Wooldridge (1999)

La forma de proceder es la siguiente: supongamos que eliminamos el periodo T, dejando las ecuaciones

$$\begin{aligned}\ddot{y}_{i1} &= \ddot{x}_{i1}\beta + \ddot{u}_{i1} \\ &\dots\dots\dots \\ \ddot{y}_{i,T-1} &= \ddot{x}_{i,T-1}\beta + \ddot{u}_{i,T-1}\end{aligned}$$

la diferencia con la ecuación  $\ddot{y}_i = \ddot{x}_i\beta + \ddot{u}_i$  esta en los distintos rangos de los vectores y matrices; así

$$\ddot{y}_i \text{ es } (T-1) \times 1, \quad \ddot{x}_i \text{ es } (T-1) \times K \quad \text{y} \quad \ddot{u}_i \text{ es } (T-1) \times 1.$$

Definimos la matriz definida positiva  $(T-1) \times (T-1) \quad \Omega \equiv E[\ddot{u}_i\ddot{u}_i']$ . No necesitamos explicitar la dependencia de  $\Omega$  sobre  $Q_T$  y  $\wedge$ ; el punto clave es que, si no hacemos restricciones sobre  $\wedge$ , entonces  $\Omega$  será también no restringido.

Para estimar  $\Omega$ , estimamos  $\beta$  por efectos fijos, después eliminamos el último periodo de cada  $i$  y definimos los residuos  $(T-1) \times 1$  como  $\hat{\ddot{u}}_i = \ddot{Y}_i - \ddot{X}_i\hat{\beta}_{FE} \quad i = 1, 2, \dots, N$ .

$$\text{Un estimador consistente de } \Omega \text{ es } \hat{\Omega} = \frac{\sum_{i=1}^N \hat{\ddot{u}}_i \hat{\ddot{u}}_i'}{N}$$

El estimador de efectos fijos por MCG (EFMCG) se define por:

$$\hat{\beta}_{EFMCG} = \left( \sum_{i=1}^N \ddot{X}_i' \hat{\Omega}^{-1} \ddot{X}_i \right)^{-1} \left( \sum_{i=1}^N \ddot{X}_i' \hat{\Omega}^{-1} \ddot{Y}_i \right) \text{ donde } \ddot{X}_i \text{ y } \ddot{Y}_i \text{ están definidas con el último periodo eliminado.}$$

Para que el estimador EFMCG sea consistente debemos reemplazar la hipótesis EF.2 por una nueva condición de rango.

Hipótesis EFMCG.2:  $\text{rango } E[\ddot{X}_i' \Omega^{-1} \ddot{X}_i] = K$

Con las hipótesis EF.1 y EFMCG.2, el estimador EFMCG es un estimador consistente. Cuando añadimos la hipótesis EFMCG.3, la varianza asintótica es fácil de estimar:

$$A \text{ var}(\hat{\beta}_{FEGLS}) = \left( \sum_{i=1}^N \ddot{X}_i' \Omega^{-1} \ddot{X}_i \right)^{-1}$$

La suma de los residuos al cuadrado de MCG puede usarse para testar múltiples restricciones. Hay que recordar que  $G = T-1$  en el estadístico F de la ecuación.

El estimador EFMCG fue propuesto por Kiefer (1980) para el caso de que  $C_i$  sean tratados como parámetros. El procedimiento permite estimar  $\beta$  de forma consistente cuando  $C_i$  es visto como aleatorio y pudiera estar correlacionado arbitrariamente con  $x_{it}$

### 3.1.8.- MÉTODO DE PRIMERAS DIFERENCIAS.

#### 3.1.8.1.- INFERENCIA.

Vamos a estudiar la transformación en diferencias en el caso general del modelo  $y_{it} = x_{it}\beta + c_i + u_{it} \quad t=1,2,\dots,T$ .

Las hipótesis de partida son:

Hipótesis PD.1:  $E[u_{it} / x_i, c_i] = 0 \quad t=1,2,\dots,T$  que es la misma hipótesis EF.1

El modelo y la interpretación de los coeficientes  $\beta$  es la misma que el método de Efectos Fijos. Para poder aplicar este método de primeras diferencias trabajamos de la forma siguiente:

Tomamos nuestro modelo  $y_{it} = x_{it}\beta + c_i + u_{it} \quad t = 1, 2, \dots, T$  y lo retardamos un periodo y a continuación restamos uno de otro llegando a la expresión:

$$\Delta y_{it} = \Delta x_{it}\beta + \Delta u_{it} \quad t = 2, 3, \dots, T$$

Donde  $\Delta y_{it} = y_{it} - y_{i,t-1}$ ,  $\Delta x_{it} = x_{it} - x_{i,t-1}$  y por último  $\Delta u_{it} = u_{it} - u_{i,t-1}$ . El efecto de esta transformación es eliminar el efecto no observado  $c_i$ . Al tomar diferencias perdemos el primer periodo de cada sección cruzada  $i$ , de manera que ahora tenemos  $T-1$  periodos temporales para cada  $i$  en lugar de los  $T$  que teníamos antes.

La ecuación  $\Delta y_{it} = \Delta x_{it}\beta + \Delta u_{it} \quad t = 2, 3, \dots, T$  pone de manifiesto que los elementos a incluir en  $x_{it}$  deben variar en el tiempo, en caso contrario serán eliminados en la transformación. Una ventaja añadida a este método es que si bien el intercepto en la ecuación original elimina las diferencias, la ecuación  $\Delta y_{it} = \Delta x_{it}\beta + \Delta u_{it} \quad t = 2, 3, \dots, T$  puede contener cambios en la variables ficticias temporales si  $x_{it}$  contiene ficticias temporales.

El estimador de primeras diferencias (PD),  $\hat{\beta}_{PD}$  es el estimador MCO Agrupados de la regresión  $\Delta y_{it}$  sobre  $\Delta x_{it} \quad t = 2, 3, \dots, T, \quad i = 1, 2, \dots, N$

Bajo la hipótesis PD.1, la estimación MCO Agrupados de la ecuación en primeras diferencias será consistente si se cumple

$$E[\Delta x'_{it} \Delta u_{it}] = 0 \quad t = 2, 3, \dots, T,$$

Para ello es preciso que se cumpla que  $E[\Delta u_{it} / \Delta x_{i2}, \Delta x_{i3}, \dots, \Delta x_{iT}] = 0 \quad t = 2, 3, \dots, T$  que significa que el estimador PD es condicionalmente insesgado en  $x$ .

Para llegar a la hipótesis  $E[\Delta x'_{it} \Delta u_{it}] = 0 \quad t = 2, 3, \dots, T$  tenemos que tomar una hipótesis menos estricta que la PD.1. El punto clave es que la hipótesis  $E[\Delta x'_{it} \Delta u_{it}] = 0 \quad t = 2, 3, \dots, T$  falla si  $u_{it}$  está correlacionado con  $x_{i,t-1}, x_{it}$  o  $x_{i,t+1}$ ; por lo tanto asumiremos que  $x_{it}$  está incorrelacionado con  $u_{it}$  para todo  $t$  y  $s$ .

Para completar nuestro conjunto de hipótesis de partida, establecemos la condición de rango para el estimador PD.

Hipótesis PD.2 
$$\text{rango} \left[ \sum_{t=2}^T E[\Delta x'_{it} \Delta u_{it}] \right] = K$$

En la práctica esta hipótesis supone eliminar a las variables explicativas constantes en el tiempo y a las variables que varían en el tiempo con colinealidad perfecta.

Asumiendo que los datos han sido ordenados como decíamos anteriormente, el método de Primeras Diferencias es fácil de implementar teniendo cuidado de qué observaciones transformadas son válidas y cuáles no. Las diferencias para las observaciones  $1, T+1, 2T+1, 3T+1, \dots$  y  $(N-1)T+1$  deben establecerse como “perdidas”. Dichas observaciones corresponden a los primeros periodos de cada unidad de sección cruzada en el conjunto de datos originales y por definición no hay primera diferencia para la observación  $t = 1$ . Hay que tener cuidado con las diferencias en el primer periodo temporal de la unidad  $i + 1$  y el último periodo de la unidad  $i$ .

Una razón para preferir el estimador PD al estimador EF es que PD es más fácil de obtener sin necesidad de ningún software especial. Pero aparte de esta razón de sencillez de implementación, ¿existe algunas razones estadísticas para preferir PD a EF?

La respuesta la encontramos en Wooldridge (2002:280), bajo las hipótesis EF.1 – EF.3 el estimador de Efectos Fijos es asintóticamente eficiente dentro de la clase de estimadores que utilizan la hipótesis de exogeneidad estricta EF.1.

Por tanto el estimador de Primera Diferencia es menos eficiente que el estimador de Efectos Fijos bajo las hipótesis EF.1 – EF.3. La hipótesis EF.3 es la clave de la eficiencia de EF. Esta hipótesis asume homoscedasticidad y no autocorrelación en  $u_{it}$ . La asunción de que los  $\{u_{it} : t=1,2,\dots,T\}$  estén no autocorrelacionados puede resultar demasiado estricta. Una alternativa a dicha hipótesis es suponer que la primera diferencia del error idiosincrásico  $\{e_{it} = \Delta u_{it}, t=2,3,\dots,T\}$  esté no autocorrelacionado (y tenga varianza constante)

Hipótesis PD.3 
$$E[e_i e_i' / x_{i1} \dots x_{iT}, c_i] = \sigma_e^2 I_{T-1}$$

Donde  $e_i$  es el vector  $(T-1) \times 1$  que contiene  $e_{it}$ ,  $t=2,3,\dots,T$

Bajo la hipótesis PD.3 podemos escribir  $u_{it} = u_{i,t-1} + e_{it}$  por lo tanto, la no autocorrelación en  $e_{it}$  implica que  $u_{it}$  es un paseo aleatorio. Un paseo aleatorio tiene una dependencia temporal considerable y por lo tanto la hipótesis PD.3 representa el extremo opuesto a la hipótesis EF.3.

Bajo las hipótesis PD.1 – PD.3 puede demostrarse que el estimador PD es más eficiente dentro de la clase de estimadores que utilizan la hipótesis EF. 1 de exogeneidad estricta.

El estimador asintótico de la varianza del estimador lo obtenemos a partir del estimador de MCO Agrupados.

$$\hat{Avar}(\hat{\beta}_{PD}) = \hat{\sigma}_e^2 (\Delta X' \Delta X)^{-1}$$

donde  $\hat{\sigma}_e^2$  es un estimador consistente de  $\sigma_e^2$ . El estimador más simple de cálculo se obtiene con los residuos MCO

$$\hat{e}_{it} = \Delta y_{it} - \Delta x_{it} \hat{\beta}_{PD}$$

de la regresión  $\Delta y_{it}$  sobre  $\Delta x_{it}$   $t = 2, 3, \dots, T$ ,  $i = 1, 2, \dots, N$ . Un estimador consistente de  $\sigma_e^2$  es:

$$\hat{\sigma}_e^2 = \frac{\sum_{i=1}^N \sum_{t=2}^T \hat{e}_{it}^2}{N(T-1) - K}$$

que es el estimador tradicional que se utiliza en la regresión  $\Delta y_{it}$  sobre  $\Delta x_{it}$   $t = 2, 3, \dots, T$ ,  $i = 1, 2, \dots, N$ . Estas ecuaciones muestran que, bajo las hipótesis PD.1 – PD.3, los errores MCO habituales de la regresión de primeras diferencias son asintóticamente válidos.

Además, bajo la hipótesis PD.3, todos los estadísticos obtenidos de la regresión agrupada de los datos de primeras diferencias son asintóticamente válidos, incluyendo los estadísticos F basados en suma de cuadrados de residuos.

### 3.1.8.2.- MATRIZ DE VARIANZA ROBUSTA.

Si la hipótesis PD.3 no puede mantenerse, nuevamente Wooldridge (2002: 282) nos dice que podemos calcular una matriz de varianzas robusta. El estimador a utilizar en este contexto es:

$$\hat{Avar}(\hat{\beta}_{PD}) = (\Delta x' \Delta x)^{-1} \left( \sum_{i=1}^N \Delta x_i' \hat{e}_i \hat{e}_i' \Delta x_i \right) (\Delta x' \Delta x)^{-1}$$

donde  $\Delta x$  denota la matriz  $N(T-1) \times K$  de primeras diferencias.



### 3.1.8.3.- CONTRASTE DE AUTOCORRELACIÓN.

Bajo la hipótesis PD.3, los errores  $e_{it} = \Delta u_{it}$  deberían estar no autocorrelacionados. Esta hipótesis es posible contrastarla con los residuos de la regresión  $\Delta y_{it}$  sobre  $\Delta x_{it}$   $t=2,3,\dots,T$ ,  $i=1,2,\dots,N$ . Como la hipótesis de exogeneidad estricta se mantiene, Wooldridge (2002:282) nos recomienda aplicar el contraste siguiente:

$$\hat{e}_{it} = \hat{\rho}_1 \hat{e}_{i,t-1} + error_{it} \quad t=3,4,\dots,T \quad i=1,2,\dots,N$$

Este contraste es el test habitual  $t$  sobre  $\hat{\rho}_1$ . Este test se basa en una regresión con  $T-2$  periodos temporales. Cuando  $T=2$  el test es imposible de realizar y por último, con  $T=3$  la regresión, se convierte en una regresión de sección cruzada ya que perdemos los periodos  $T=1$  y  $T=2$ .

Si el error idiosincrásico  $\{u_{it} : t=1,2,\dots,T\}$  está incorrelacionado,  $\{e_{it} : t=2,3,\dots,T\}$  estará autocorrelacionado. Bajo la hipótesis PD.3 es posible demostrar que  $Corr(e_{it}, e_{i,t-1}) = -0,5$ . En cualquier caso, encontrar una correlación temporal importante en los  $e_{it}$  es una manera de asegurarnos de que debemos calcular una matriz de varianza robusta para el estimador PD.

### 3.1.9.- COMPARACIÓN DE ESTIMADORES.

#### 3.1.9.1.- EFECTOS FIJOS VERSUS PRIMERAS DIFERENCIAS.

Cuando tenemos sólo dos periodos, la estimación de efectos fijos y primeras diferencias producen idénticas estimaciones e inferencia.

El método de primeras diferencias es más fácil de implementar y todos los procedimientos que suelen aplicarse en los modelos de sección cruzada, son susceptibles de aplicarse directamente.

Cuando  $T > 2$  la elección entre ambos estimadores depende de las hipótesis a cerca del error idiosincrásico  $u_{it}$ . El estimador EF es más eficiente bajo la hipótesis EF.3 (los  $u_{it}$  no están autocorrelacionados), mientras que el estimador PD es más eficiente cuando  $u_{it}$  siguen un paseo aleatorio.

Si los estimadores EF y PD difieren tanto que su diferencia no puede ser atribuida al error muestral, entonces, tal y como nos muestra Wooldridge (2002: 284) deberíamos preocuparnos por la hipótesis de EXOGENEIDAD ESTRICTA.

Si los  $u_{it}$  están correlacionados con las  $x_{is}$  para cualquier  $t$  y  $s$ , EF y PD tendrán distintos límites en probabilidad. Problemas de endogeneidad habituales tales como errores de medida, variables temporales omitidas, etc... provocan correlación entre  $x_{it}$  y  $u_{it}$ , esto es, correlación contemporánea, lo que conduce a que ambos estimadores EF y PD sean NO consistentes y tengan límites en probabilidad distintos.

Además, la correlación entre  $u_{it}$  y  $x_{is}$  para  $s \neq t$  hace que de nuevo los estimadores EF y PD sean inconsistentes.

La forma de contrastar formalmente las hipótesis de consistencia de los estimadores EF y PD es a través de un test de Hausman (1978). Contraste que permite comparar los estimadores EF y PD.

Cuando  $T = 2$  la forma de testar exogeneidad estricta es utilizar la ecuación  $\Delta y_i = \Delta x_i \beta + \Delta u_i$  e introducir en dicha ecuación, como variables explicativas adicionales  $x_{i1}$  y  $x_{i2}$ . Estas variables adicionales no deberían ser significativas en la ecuación de primeras diferencias.

Para más de dos periodos temporales, en un test de exogeneidad estricta se trataría de contrastar la hipótesis nula  $H_0: \gamma = 0$  en la ecuación  $\Delta y_t = \Delta x_t \beta + w_t \gamma + \Delta u_t \quad t = 2, 3 \dots T$  en la que  $w_t$  es un subconjunto de  $x_t$  (en el que están excluidas las variables ficticias temporales). El estadístico habitual F continúa siendo válido asintóticamente.

Con efectos fijos, un test de exogeneidad estricta en el caso de que  $T > 2$  se obtiene especificando la ecuación  $y_{it} = x_{it} \beta + w_{i,t+1} \delta + c_i + u_{it} \quad t = 1, 2 \dots T-1$  en la que  $w_{i,t+1}$  es de nuevo un subconjunto de  $x_{i,t+1}$ . Bajo exogeneidad estricta,  $\delta = 0$  y podemos implementar el test utilizando la estimación de efectos fijos

### 3.1.9.2.- RELACIÓN ENTRE LOS ESTIMADORES DE EFECTOS ALEATORIOS Y EFECTOS FIJOS.

Cuando las variables claves en  $x_i$  no varían mucho en el tiempo, los métodos de efectos fijos y primera diferencia pueden llevarnos a estimaciones imprecisas. Entonces resulta una buena idea utilizar estimación de efectos aleatorios para conocer algo más acerca de los parámetros poblacionales.

Cuando el análisis de efectos aleatorios es apropiado; esto es, cuando  $c_i$  es ortogonal a  $x_{it}$  (no hay relación entre ellos) entonces los estimadores de efectos aleatorios pueden tener varianzas más pequeñas que los estimadores de efectos fijos o de primera diferencia. Vamos a obtener una expresión de los estimadores EA que nos permitirá compararlos con los estimadores de EF. Para ello seguiremos a Wooldridge (2002:286)

Escribimos  $\Omega$  bajo la estructura de efectos aleatorios de la forma siguiente:

$$\Omega = \sigma_u^2 I_T + \sigma_c^2 J_T J_T'$$

para utilizarlo en esta expresión conocemos:

$$\begin{aligned} J_T J_T' &= T \\ P_T &= J_T (J_T' J_T)^{-1} J_T' \\ P_T &= I_T - Q_T \quad I_T = P_T + Q_T \quad Q_T = I_T - P_T \\ Q_T &= I_T - J_T (J_T' J_T)^{-1} J_T' \end{aligned}$$

En la expresión original, sustituimos  $I_T = P_T + Q_T$  y  $J_T J_T' = T$

$$\begin{aligned} \Omega &= \sigma_u^2 I_T + \sigma_c^2 J_T J_T'; \quad \Omega = \sigma_u^2 (P_T + Q_T) + T \sigma_c^2 J_T (J_T' J_T)^{-1} J_T'; \\ \Omega &= \sigma_u^2 P_T + \sigma_u^2 Q_T + T \sigma_c^2 \end{aligned}$$

A continuación multiplicamos y dividimos  $Q_T$  por  $\sigma_u^2 + T\sigma_c^2$

$$\Omega = (\sigma_u^2 + T\sigma_c^2)P_T + (\sigma_u^2 + T\sigma_c^2) \frac{\sigma_u^2}{(\sigma_u^2 + T\sigma_c^2)} Q_T \quad \text{llamamos } n = \frac{\sigma_u^2}{(\sigma_u^2 + T\sigma_c^2)}$$

y sustituimos

$$\Omega = (\sigma_u^2 + T\sigma_c^2)P_T + (\sigma_u^2 + T\sigma_c^2)nQ_T;$$

reagrupamos

$$\Omega = \sigma_u^2 P_T + T\sigma_c^2 P_T + n\sigma_u^2 Q_T + nT\sigma_c^2 Q_T$$

$$\Omega = (\sigma_u^2 + T\sigma_c^2)(P_T + nQ_T)$$

A continuación definimos  $S_T = P_T + nQ_T$ , Entonces  $S_T^{-1} = P_T + \left(\frac{1}{n}\right)Q_T$ . Por otra parte

calculamos  $S_T^{-1/2} = P_T + \left(\frac{1}{\sqrt{n}}\right)Q_T$  ya que multiplicando esta matriz por si misma

obtenemos  $S_T^{-1}$  (la matriz es simétrica dado que  $P_T$  y  $Q_T$  lo son),

Ahora vamos a calcular una expresión que nos permita pasar de un tipo de estimación a otra de la forma más sencilla; dicho en otras palabras, relacionar como procedentes de un solo caso la estimación Efectos Fijos, Efectos Aleatorios y MCO Agrupados. Para ello procedemos de la siguiente manera.

$$S_T^{-1/2} = P_T + \left(\frac{1}{\sqrt{n}}\right)Q_T \quad ; \quad S_T^{-1/2} = \frac{1}{\sqrt{n}}(Q_T + \sqrt{n}P_T) \quad ; \quad S_T^{-1/2} = (\sqrt{n})^{-1}[(I_T - P_T) + \sqrt{n}P_T]$$

$$S_T^{-1/2} = (\sqrt{n})^{-1}[I_T - (1 - \sqrt{n})P_T] \quad , \text{ sumando y restando 1 a la primera parte tenemos}$$

$$S_T^{-1/2} = (1 - (1 - \sqrt{n}))^{-1}[I_T - (1 - \sqrt{n})P_T] \quad , \text{ identificando a } \lambda = (1 - \sqrt{n}) \text{ y sustituyendo tenemos}$$

$$S_T^{-1/2} = (1 - \lambda)^{-1}[I_T - \lambda P_T]$$

Por otra lado, partiendo de que  $\Omega = (\sigma_u^2 + T\sigma_c^2)(P_T + nQ_T)$ ;  $\Omega = (\sigma_u^2 + T\sigma_c^2)S_T$

queremos calcular  $\Omega^{-1/2} = (\sigma_u^2 + T\sigma_c^2)^{-1/2} S_T^{-1/2}$  sustituyendo en este la expresión

obtenida anteriormente para  $S_T^{-1/2}$  tenemos  $\Omega^{-1/2} = (\sigma_u^2 + T\sigma_c^2)^{-1/2} (1 - \lambda)^{-1}[I_T - \lambda P_T]$

Ahora teniendo en cuenta que  $\lambda = (1 - \sqrt{n})$  y sustituyéndolo en la expresión anterior llegamos a

$$\Omega^{-1/2} = \frac{1}{(\sigma_u^2 + T\sigma_c^2)^{1/2}} \cdot \frac{1}{(1-\lambda)} \cdot [I_T - \lambda P_T] \quad ; \quad \Omega^{-1/2} = \frac{1}{(\sigma_u^2 + T\sigma_c^2)^{1/2}} \cdot \frac{1}{(1-1+\sqrt{n})} \cdot [I_T - \lambda P_T] \quad ;$$

$$\Omega^{-1/2} = \frac{1}{(\sigma_u^2 + T\sigma_c^2)^{1/2}} \cdot \frac{1}{(\sqrt{n})} \cdot [I_T - \lambda P_T]$$

a continuación tenemos en cuenta que  $n = \frac{\sigma_u^2}{(\sigma_u^2 + T\sigma_c^2)}$  y lo sustituimos en la expresión anterior obteniendo:

$$\Omega^{-1/2} = \frac{1}{(\sigma_u^2 + T\sigma_c^2)^{1/2}} \cdot \frac{1}{\left( \frac{\sigma_u^2}{(\sigma_u^2 + T\sigma_c^2)} \right)^{1/2}} \cdot [I_T - \lambda P_T] \quad ; \quad \Omega^{-1/2} = \frac{1}{\sigma_u} \cdot [I_T - \lambda P_T]$$

Asumimos por el momento que conocemos  $\lambda$ . Entonces el estimador EA se obtiene estimando la ecuación transformada  $C_T Y_i = C_T X_i \beta + C_T v_i$  por MCP, donde  $C_T = [I_T - \lambda P_T]$ . Escribimos la ecuación transformada como  $\tilde{y}_i = \tilde{X}_i \beta + \tilde{v}_i$ .

La varianza de la matriz  $\tilde{v}_i$  es  $E[\tilde{v}_i \tilde{v}_i'] = C_T \Omega C_T = \sigma_u^2$ , que verifica que  $\tilde{v}_i$  tiene la matriz de varianza ideal para el sistema de estimación MCO.

El elemento  $t$  de  $\tilde{y}_i$  es  $y_{it} - \lambda \bar{y}_i$  y lo mismo para  $\tilde{X}_i$ . Por lo tanto, la estimación por el sistema MCO de la ecuación  $\tilde{y}_i = \tilde{X}_i \beta + \tilde{v}_i$  es justamente la estimación MCO Agrupada de:

$$y_{it} - \lambda \bar{y}_i = (x_{it} - \lambda \bar{x}_i) \beta + (v_{it} - \lambda \bar{v}_i) \text{ para todo } t \text{ e } i.$$

Los errores en esta ecuación son homoscedásticos y no están autocorrelacionados bajo la hipótesis EA.3; por lo tanto, satisfacen las condiciones claves para el análisis de MCO Agrupados.

El estimador EA factible reemplaza el desconocido  $\lambda$  por su estimador  $\hat{\lambda}$ , por lo que  $\hat{\beta}_{EA}$  puede calcularse de la regresión de MCO Agrupados.

$$\tilde{y}_{it} \text{ sobre } \tilde{x}_{it} \quad t=1,2,\dots,T, \quad i=1,2,\dots,N$$

Donde ahora  $\tilde{x}_{it} = x_{it} - \hat{\lambda}\bar{x}_i$  y  $\tilde{y}_{it} = y_{it} - \hat{\lambda}\bar{y}_i$  para todo  $t$  e  $y$ . Por tanto el estimador de efectos aleatorios puede escribirse como

$$\hat{\beta}_{EA} = \left( \sum_{i=1}^N \sum_{t=1}^T \tilde{x}_{it}' \tilde{x}_{it} \right)^{-1} \left( \sum_{i=1}^N \sum_{t=1}^T \tilde{x}_{it}' \tilde{y}_{it} \right)$$

La estimación habitual de la varianza en la regresión de MCO Agrupados de  $\tilde{y}_{it}$  sobre  $\tilde{x}_{it}$  es  $\frac{SSR}{NT-K}$  que resulta ser un estimador consistente de  $\sigma_u^2$ . Los estadísticos habituales  $t$  y  $F$  de la regresión de MCO Agrupados son asintóticamente válidos bajo las hipótesis EA.1 – EA.3.

Esta última expresión del estimador de efectos aleatorios se obtiene por una “QUASITIME DEMEANING”. Esto es, el lugar de eliminar totalmente el promedio temporal en las variables explicativas y en la variable dependiente, los efectos aleatorios eliminan una fracción de dicho promedio temporal.

Si  $\hat{\lambda}$  está cerca de la unidad, las estimaciones de efectos aleatorios y las de efectos fijos tienden a estar cerca. Para verlo escribimos  $\hat{\lambda}$  como:

$$\hat{\lambda} = 1 - \left[ \frac{1}{1 + T \left( \frac{\hat{\sigma}_c^2}{\hat{\sigma}_u^2} \right)} \right]^{1/2}$$

Donde  $\hat{\sigma}_u^2$  y  $\hat{\sigma}_c^2$  son los estimadores consistentes de  $\sigma_u^2$  y  $\sigma_c^2$  esto es:

$\hat{\sigma}_u^2 = \hat{\sigma}_v^2 - \hat{\sigma}_c^2$ ,  $\hat{\sigma}_v^2 = \frac{1}{(NT-K)} \sum_{i=1}^N \sum_{t=1}^T \hat{v}_{it}^2$ , siendo  $\hat{v}_{it}$  los residuos de la regresión de MCO

Agrupados y  $\hat{\sigma}_c^2 = \frac{1}{\left[ N \frac{T(T-1)}{2} - K \right]} \sum_{i=1}^N \sum_{t=1}^{T-1} \sum_{s=t+1}^T \hat{v}_{it} \hat{v}_{is}$ . Cuando  $T \left( \frac{\hat{\sigma}_c^2}{\hat{\sigma}_u^2} \right)$  es grande, el

segundo término de  $\hat{\lambda}$  es pequeño y en estos casos  $\hat{\lambda}$  está cerca de la unidad.

En efecto,  $\hat{\lambda} \rightarrow 1$  cuando  $T \rightarrow \infty$  o cuando  $\left( \frac{\hat{\sigma}_c^2}{\hat{\sigma}_u^2} \right) \rightarrow \infty$ .

Para una  $T$  grande no es sorprendente encontrar estimaciones similares de efectos fijos y de efectos aleatorios. Incluso con  $T$  pequeña los efectos aleatorios pueden estar cerca de los efectos fijos si la estimación de la varianza de  $c_i$  es relativamente grande respecto a la varianza estimada de  $u_{it}$ . Cuando  $\lambda$  se aproxima a la unidad, la precisión del estimador de efectos aleatorios se aproxima al estimador de efectos fijos, y los efectos de las variables explicativas constantes en el tiempo se vuelven difíciles de estimar.

Las expresiones  $\hat{\beta}_{EA} = \left( \sum_{i=1}^N \sum_{t=1}^T \tilde{x}_{it}' \tilde{x}_{it} \right)^{-1} \left( \sum_{i=1}^N \sum_{t=1}^T \tilde{x}_{it}' \tilde{y}_{it} \right)$  y  $\hat{\lambda} = 1 - \left[ \frac{1}{1 + T \left( \frac{\hat{\sigma}_c^2}{\hat{\sigma}_u^2} \right)} \right]^{\frac{1}{2}}$  muestran

también la relación existente entre los efectos aleatorios y los MCO Agrupados. Los MCO Agrupados se obtienen haciendo  $\hat{\lambda} = 0$ , que nunca es exactamente cierto pero podría estar cerca. En la práctica  $\hat{\lambda}$  no está habitualmente cerca de 0 porque este resultado requeriría que  $\hat{\sigma}_u^2$  fuera relativamente grande respecto a  $\hat{\sigma}_c^2$ .

La regresión  $\tilde{y}_{it}$  sobre  $\tilde{x}_{it}$  utilizando datos “quasi-time demeaned” tiene bastantes otros usos prácticos. Dado que es una regresión de MCO Agrupados en la que es asintóticamente lo mismo usar  $\lambda$  en lugar de  $\hat{\lambda}$ , podemos obtener errores estándar que sean robustos a la heteroscedasticidad arbitraria en  $c_i$  y  $u_{it}$ , así como la autocorrelación arbitraria en  $\{u_{it}\}$ . Para ello necesitamos calcular los errores estándar robustos  $t$  y los estadísticos  $F$  para la regresión MCO Agrupados. Con los



residuos de la regresión  $\tilde{y}_{it}$  sobre  $\tilde{x}_{it}$  a los que llamamos,  $\hat{r}_{it}$  se utilizan para contrastar la correlación serial en  $r_{it} = v_{it} - \lambda \bar{v}_i$  que están incorrelacionados en el tiempo bajo la hipótesis EA.3a. Si detectamos correlación en el tiempo en  $r_{it}$ , concluimos que la hipótesis EA.3a es falsa y este resultado significa que los  $u_{it}$  están autocorrelacionados.

### 3.1.9.3.- TEST DE HAUSMAN PARA COMPARAR ESTIMADORES DE EFECTOS ALEATORIOS Y EFECTOS FIJOS.

Dado que el elemento clave para elegir entre los estimadores de Efectos Fijos y los de Efectos Aleatorios es saber si  $c_i$  y  $x_{it}$  están correlacionados, resulta básico disponer de un método para testar esta hipótesis. Hausman (1978) propuso un test basado en las diferencias entre las estimaciones de efectos aleatorios y efectos fijos.

Recordemos que los EF son consistentes cuando  $c_i$  y  $x_{it}$  están correlacionados, y los EA son inconsistentes cuando esto ocurre, por lo tanto una diferencia estadísticamente significativa se interpreta como una evidencia en contra de la hipótesis de efectos aleatorios EA.1b:  $E[c_i / x_i] = E[c_i] = 0$  donde  $x_i = (x_{i1}, x_{i2}, \dots, x_{iT})$ . Esto es, la hipótesis de que  $c_i$  y  $x_{it}$  están incorrelacionados no puede mantenerse y lo adecuado en ese caso sería utilizar el estimador EF.

Antes de obtener el test de Hausman hay dos cuestiones previas:

1. La exogeneidad estricta, la hipótesis EA.1a:  $E[u_{it} / x_i, c_i] = 0 \quad t=1, 2, \dots, T$  se mantiene bajo la hipótesis nula y la alternativa. Correlación entre  $x_{is}$  y  $u_{it}$  para cualquier  $s$  y  $t$  provoca que ambos estimadores sean inconsistentes y generalmente su límite en probabilidad difiere.

2. El contraste se implementa habitualmente asumiendo que la hipótesis EA.3:

$E[u_i u_i' / x_i, c_i] = \sigma_u^2 I_T$  y  $E[c_i^2 / x_i] = \sigma_c^2$  se mantiene bajo la hipótesis nula. Esta consideración implica que el estimador de efectos aleatorios es más eficiente que el estimador de efectos fijos y simplifica el cálculo del test estadístico. Pero es preciso advertir que la hipótesis EA.3 es una hipótesis auxiliar y no está siendo testada por el contraste de Hausman. Fallos en la hipótesis EA.2 provocan que el test de Hausman tenga un límite en distribución no estándar, lo que significa que el resultado del test podría tener asintóticamente un tamaño mayor o menor que el tamaño nominal.

Asumiendo que las hipótesis EA.1 – EA.3 se mantienen, consideramos exclusivamente el caso en el que  $x_{it}$  contiene sólo elementos que varían en el tiempo, por lo tanto los coeficientes pueden ser estimados únicamente utilizando efectos fijos. Entonces

$$A \text{ var}(\hat{\beta}_{EF}) = \sigma_u^2 \left[ E(\ddot{X}_i' \ddot{X}_i) \right]^{-1} / N$$

$$A \text{ var}(\hat{\beta}_{EA}) = \sigma_u^2 \left[ E(\tilde{X}_i' \tilde{X}_i) \right]^{-1} / N$$

Donde el término  $t$  de la columna de  $\ddot{X}_i$  es  $x_{it} - \bar{x}_i$  y el término  $t$  de la columna de  $\tilde{X}_i$  es  $x_{it} - \lambda \bar{x}_i$ .

$$E[\tilde{X}_i' \tilde{X}_i] - E[\ddot{X}_i' \ddot{X}_i] = E[X_i'(I_T - \lambda P_T) X_i] - E[X_i'(I_T - P_T) X_i] =$$

$$= (1 - \lambda) E[X_i' P_T X_i] = (1 - \lambda) T E[\bar{X}_i' \bar{X}_i]$$

De lo que se sigue que  $\left[ A \text{ var}(\hat{\beta}_{EA}) \right]^{-1} - \left[ A \text{ var}(\hat{\beta}_{EF}) \right]^{-1}$  es definida positiva, implicando que  $A \text{ var}(\hat{\beta}_{EF}) - A \text{ var}(\hat{\beta}_{EA})$  es también definida positiva. Dado que  $\lambda \rightarrow 1$  cuando  $T \rightarrow \infty$ , esta expresión muestra que la varianza asintótica del estimador de Efectos Aleatorios tiende a la del estimador de Efectos Fijos cuando  $T$  es grande.

En su forma original el test de Hausman se calcula de la siguiente manera. Sea  $\hat{\delta}_{EA}$  el vector de estimaciones de efectos aleatorios, sin los coeficientes de las variables constantes en el tiempo o variables temporales agregadas, y sea  $\hat{\delta}_{EF}$  el correspondiente a las estimaciones de efectos fijos; sea cada uno un vector de  $M \times 1$ . Entonces

$$H = (\hat{\delta}_{EF} - \hat{\delta}_{EA})' \left[ A \text{var}(\hat{\delta}_{EF}) - A \text{var}(\hat{\delta}_{EA}) \right]^{-1} (\hat{\delta}_{EF} - \hat{\delta}_{EA})$$

Que se distribuye asintóticamente como una  $\chi^2_M$  bajo las hipótesis EA.1 a EA.3.

El estimador habitual de  $\left[ A \text{var}(\hat{\delta}_{EA}) \right]$  y  $\left[ A \text{var}(\hat{\delta}_{EF}) \right]$  es el que suele utilizarse en la expresión del test de Hausman, pero si usamos diferentes estimaciones de  $\sigma_u^2$ , la matriz  $\left[ A \text{var}(\hat{\delta}_{EF}) - A \text{var}(\hat{\delta}_{EA}) \right]$  no necesita ser definida positiva. Por lo tanto, la forma más sencilla de implementar el contraste de Hausman es utilizar la estimación de efectos fijos o de efectos aleatorios de  $\sigma_u^2$  en ambos lugares.

Otras veces, podemos estar interesados en un sólo parámetro, en ese caso utilizamos un estadístico  $t$  que ignora los otros parámetros. Sea  $\delta$  el elemento de  $\beta$  que deseamos contrastar. El test de Hausman se calcula con su versión para el estadístico  $t$

$$\frac{(\hat{\delta}_{EF} - \hat{\delta}_{EA})}{\left[ \left( se(\hat{\delta}_{EF}) \right)^2 - \left( se(\hat{\delta}_{EA}) \right)^2 \right]^{1/2}} \text{ donde los errores estándar se calculan bajo las hipótesis}$$

habituales. Bajo las hipótesis EA.1 a EA.3, el estadístico  $t$  sigue una distribución asintóticamente normal.

Para testar más de un parámetro, es más sencillo usar una versión distinta del estadístico F del test de Hausman.

Sean  $\tilde{x}_{it}$  y  $\tilde{y}_{it}$  los datos “*quasi-demeaned*” definidos previamente;  $w_{it}$  que denota un subconjunto de  $1 \times M$  elementos de  $x_{it}$  variables que varían en el tiempo (excluimos las variables ficticias temporales); en este subconjunto se pueden incluir todos los elementos de  $x_{it}$  que varíen en  $i$ , en  $t$  o un subconjunto.

$\ddot{w}_{it}$  es la versión “*time-demeaned*” de  $w_{it}$  y consideramos el modelo extendido.

$$\tilde{y}_{it} = \tilde{x}_{it}\beta + \ddot{w}_{it}\xi + error_{it} \quad t=1,\dots,T \quad i=1,\dots,N$$

Donde  $\xi$  es un vector  $M \times 1$ . El término de error es complicado porque  $\hat{\lambda}$  reemplaza a  $\lambda$  en la obtención de los datos “*quasi-demeaned*”, pero deben tratarse como homoscedásticos y no autocorrelacionados por el hecho que al sustituir a  $\lambda$  por  $\hat{\lambda}$  no importa asintóticamente.

Ahora el test de Hausman puede implementarse para testar  $H_o: \xi=0$  utilizando un análisis de MCO Agrupados. La aproximación más sencilla es calcular el estadístico F. La SSR restringida se obtiene de la regresión de MCO Agrupados utilizada para obtener  $\hat{\beta}_{EA}$ :

$\tilde{y}_{it}$  sobre  $\tilde{x}_{it} \quad t=1,2\dots T, \quad i=1,2\dots N$ , siendo

$$\tilde{x}_{it} = x_{it} - \hat{\lambda}\bar{x}_i \quad y \quad \tilde{y}_{it} = y_{it} - \hat{\lambda}\bar{y}_i \quad para \quad todo \quad t \text{ e } y.$$

Llamamos a esta suma de cuadrados de residuos  $SSR_r$ .

La SSR no restringida viene de la regresión de MCO Agrupados de  $\tilde{y}_{it} = \tilde{x}_{it}\beta + \ddot{w}_{it}\xi + error_{it} \quad t=1,\dots,T \quad i=1,\dots,N$ . Entonces el estadístico F es

$$F = \frac{(SSR_r - SSR_{ur})}{SSR_{ur}} \cdot \frac{(NT - K - M)}{M}$$

Bajo  $H_0$  (con las hipótesis EA.1 a EA.3)  $F$  puede tratarse como una  $F_{M, NT-K-M}$  grados de libertad  $(M \cdot F \xrightarrow{a} \chi_M^2)$ .

Si la hipótesis EA.3 no se cumple, entonces necesitaríamos usar una forma robusta del test de Hausman. Probablemente la aproximación más simple es testar  $H_o: \xi = 0$  vía un estadístico robusto de Wald en el contexto de la estimación de MCO Agrupados de  $\tilde{y}_{it} = \tilde{x}_{it}\beta + \ddot{w}_{it}\xi + error_{it}$  o con  $\bar{w}_i$  en lugar de  $\ddot{w}_{it}$ . El test robusto debería tener en cuenta la autocorrelación en el tiempo y la heteroscedasticidad general.

#### **4.- CONTRASTES EN DATOS DE PANEL.**

El objetivo que se persigue en este capítulo es pasar revista a los contrastes que más adelante se van a ir necesitando; esto es, dotarnos del cuerpo teórico que permitirá ir resolviendo los posibles problemas que puedan presentarse.

Ni que decir tiene, que no se trata de una recopilación exhaustiva de todos los contrastes existentes, más bien, es una recopilación interesada, ya que se han recogido aquellos contrastes que tienen más posibilidades de utilizarse en los capítulos posteriores.

La organización del capítulo es la siguiente. En un primer momento nos centramos en el problema de la heteroscedasticidad y la autocorrelación (cómo detectarlas y cómo corregirlas), incluso se recoge un planteamiento alternativo con la utilización de estimadores robustos. Es el cuerpo principal del capítulo, dada la gran cantidad de material disponible que trata estos problemas. Posteriormente nos centramos en el contraste de Hausman, un clásico en datos de panel, que permite elegir entre los estimadores de Efectos Fijos y Aleatorios. A continuación pasamos revista a los test de Agrupabilidad, estos contrastes se utilizan para determinar si la especificación de un modelo de datos de panel tiene sentido, algo sumamente necesario cuando se pretende estimar un modelo de datos de panel. Y por último repasamos el contraste de especificación RESET. Este contraste permite determinar la existencia de no linealidades no recogidas en la estructura del modelo y por tanto incluirlas en él. El desarrollo teórico de este contraste se encuentra fundamentado en modelos de corte transversal. Aquí puede verse su aplicación en datos de panel.

Algo más a añadir, esta es la estructura de presentación en este capítulo, pero en los restantes capítulos esta estructura no se respeta, los contrastes han sido utilizados según las necesidades y no necesariamente en este orden.

## **4.1.- HETEROSCEDASTICIDAD Y AUTOCORRELACIÓN EN DATOS DE PANEL.**

De igual forma que cuando trabajamos con datos de corte transversal o con datos temporales la heteroscedasticidad y autocorrelación son problemas habituales, con datos de panel, no podía ser de otra manera y de nuevo aparecen en nuestros datos.

El esquema que vamos a seguir en este capítulo es un esquema clásico; analizaremos cada uno de los problemas por separado; estudiando la forma de detectarlos y corregirlos; exactamente igual que en los manuales tradicionales de econometría.

### **4.1.1.- HETEROSCEDASTICIDAD.**

En la perturbación aleatoria del modelo  $y_{it} = x_{it}\beta + v_{it} \quad t=1,2,\dots,T$  donde  $v_{it} = c_i + u_{it} \quad t=1,2,\dots,T$  se asume que perturbación es homoscedástica con la misma varianza en los individuos y en el tiempo. Esta hipótesis es bastante restrictiva para datos de panel, en los que las unidades de corte transversal pueden variar de tamaño y como resultado pueden exhibir variaciones diferentes.

Asumir que la perturbación aleatoria es homoscedástica cuando la heteroscedasticidad está presente producirá estimaciones consistentes de los parámetros del modelo, pero dichas estimaciones no serán eficientes. Además los errores estándar de dichas estimaciones serán sesgados a menos que se calculen correctamente errores estándar robustos ante la posible presencia de heteroscedasticidad.

#### **4.1.1.1.- DETECCIÓN DE LA HETEROSCEDASTICIDAD.**

En el capítulo anterior hemos visto algunas formas de detectar la heteroscedasticidad; entonces ¿por qué volver a tratar este tema? La razón es que hemos visto contrastes que trabajaban desde el interior de las distintas unidades de sección cruzada; esto es trabajaban con los datos de panel como si se tratara de

una unidad de sección cruzada intentando buscar la heteroscedasticidad en el interior de esa macro unidad de sección cruzada.

Los contrastes que vamos a proponer varían esencialmente en el tratamiento de la unidad de sección cruzada. Básicamente la forma de proceder será buscar la presencia de heteroscedasticidad en cada una de las unidades de sección cruzada y posteriormente comprobar si la varianza en cada una de las secciones es la misma. En este caso usamos y respetamos la existencia de las distintas secciones de sección cruzada.

Los contrastes que vamos a proponer, aunque aplicados en Estadística, no son de uso habitual en Econometría, ni tampoco en Econometría de Datos. De hecho, el primero de ellos, el contraste de Bartlett, aparece mencionado en la página 79 de la segunda edición del libro de Baltagi, *Econometric Analysis of Panel Data*, que hace referencia a la aplicación realizada por Kmenta (1986:297). Sin embargo en las ediciones posteriores, más concretamente en las ediciones 4<sup>o</sup> y 5<sup>o</sup>, este tratamiento desaparece y son sustituidos por otros contrastes diferentes. La pregunta es obvia ¿por qué mantener unos contrastes que en las ediciones posteriores de este libro desaparecen? Desde luego, resulta imposible conocer las razones que llevaron al autor a eliminar y sustituir estos contrastes, aunque pueden intuirse.

Aun así, si es lícito preguntarse por qué se han mantenido en esta tesis. El motivo radica en la tipología de datos utilizados en este trabajo. Es más, Baltagi (2001:78), cuando habla de la transformación de Fuller – Battese para heteroscedasticidad, reconoce que el procedimiento “*requiere  $T$  grande y preferiblemente  $N$  pequeño con  $T > N$* ”. Bien esta es la razón; los datos utilizados en esta tesis se caracterizan porque el número de individuos “ $N$ ” es muy pequeño, mientras que el tiempo “ $T$ ” es elevado, de tal modo que se cumple que  $T > N$ .

Esta situación no es la habitual en los datos de panel, donde normalmente se tiene un número elevado de individuos, pero durante un periodo pequeño de tiempo. Este, es en mi opinión, el motivo que llevó a Baltagi a modificar el capítulo en las ediciones siguientes de su libro y adaptarlo más a la realidad de los paneles que se suelen utilizar. Y es también la razón por la que se mantiene en esta tesis, ya que la



estructura del panel es distinta a las que habitualmente suelen encontrarse en la realidad.

La principal aportación de estos contrastes es que están diseñados para comparar la varianza entre grupos de individuos y concluir que todos tienen la misma varianza o no. Los contrastes de homoscedasticidad utilizados habitualmente en econometría, buscan el cumplimiento de la homoscedasticidad entre los individuos, por lo tanto no tienen sentido contrastes como los anteriores. O si existen individuos, como en los datos de panel, el número de observaciones disponible de cada individuo es muy limitado por lo que la aplicación práctica de estos contraste es complicada

Dicho esto, pasaremos a revisar los contrastes para detectar la presencia de heteroscedasticidad.

#### **4.1.1.1.1.- CONTRASTE DE HETEROSCEDASTICIDAD BASADO EN EL TEST DE BARTLETT**

Como ya anunciábamos unos párrafos antes, el test de Bartlett aparece mencionado en Baltagi, (2001:79), que hace una referencia a la aplicación de este contraste en Kmenta, (1986: 297).

La forma de proceder será la siguiente: En un primer paso, comprobamos la existencia de heteroscedasticidad dentro de cada una de las unidades de sección cruzada que tenemos. Recordemos que los datos existentes en cada una de ellas, son datos temporales, referidos al mismo individuo y tal y como indican los manuales de econometría, no suelen presentarse problemas de heteroscedasticidad en datos temporales referidos al mismo individuo, pero existen casos donde si es posible detectar la presencia de heteroscedasticidad.

Para comprobar la existencia de heteroscedasticidad utilizamos el test de White, en la versión propuesta por Wooldridge, (2005: 298).

4. Obtenemos los residuos al cuadrado de la regresión:  $\hat{u}_t^2$  de cada uno de los individuos
5. Calculamos el R cuadrado:  $R_c^2$  de la regresión  $\hat{u}_t^2$  sobre una constante,  $\hat{y}_t, \hat{y}_t^2 \quad t=1,2,\dots,T$  y para todos y cada uno de los individuos  $i=1,2,\dots,N$
6. El estadístico es  $nTR_c^2$  que se comporta asintóticamente como una  $\chi^2_Q$  bajo  $H_0$ .
7. Recordemos que esto lo hacemos separadamente para cada individuo, el objetivo no es otro que comprobar la presencia de heteroscedasticidad en cada uno de los individuos.

Una vez aceptada la hipótesis nula de homoscedasticidad, el siguiente paso es comprobar si las varianzas de los distintos individuos son iguales.

Para ello vamos a utilizar el test de Bartlett que aparece en NIST/SMATECH e-Handbook of Statistical Methods, (2013: 282 y 283) donde nos remiten a la formulación recogida por Snedecor y Cochran (1989).

El contraste de Bartlett se utiliza para comprobar si  $K$  muestras tienen la misma varianza; en nuestro caso si  $N$  individuos tienen la misma varianza. Este test es muy sensible a la presencia de normalidad en la muestra.

La definición es:

$$H_0: \sigma_1 = \sigma_2 = \dots = \sigma_k$$

$$H_1: \sigma_i \neq \sigma_j$$

El estadístico del test de Bartlett está diseñado para contrastar la igualdad de las varianzas entre los individuos o grupos frente a la alternativa de que las varianzas son distintas en al menos dos grupos. El estadístico es:

$$T = \frac{(N-k) \ln s_p^2 - \sum_{i=1}^k (N_i - 1) \ln s_i^2}{1 + \left( \frac{1}{3(k-1)} \right) \left( \left( \sum_{i=1}^k \frac{1}{N_i - 1} \right) - \frac{1}{N-k} \right)}$$

Donde  $s_i^2$  es la varianza del individuo o grupo  $i$ ,  $N$  es el total del tamaño muestral,  $N_i$  es el tamaño muestral del individuo o grupo  $i$ , y  $s_p^2$  es la varianza agrupada de todos los individuos o grupos. Esta varianza agrupada es un promedio ponderado de las

varianzas de los individuos o grupos y se define como: 
$$s_p^2 = \frac{\sum_{i=1}^k (N_i - 1)s_i^2}{(N - k)}$$

Para un nivel de significación  $\alpha$  rechazaremos la hipótesis nula de igualdad de varianzas cuando  $T > \chi_{k-1}^2$

#### **4.1.1.1.2.- CONTRASTE DE HETEROSCEDASTICIDAD BASADO EN EL TEST DE LEVENE**

Así como decíamos que el test de Bartlett aparece recogido en algunos manuales de econometría, no sucede lo mismo con el test de Levene, NIST/SMATECH e-Handbook of Statistical Methods, (2013: 294 y 295).

Este test, formulado originalmente por Levene (1960), presenta la ventaja respecto al anterior que es menos sensible respecto al incumplimiento de la hipótesis de partida de normalidad.

La manera de implementarlo será exactamente igual al anterior; en el primer paso buscamos el cumplimiento de la homoscedasticidad en el interior de los individuos para posteriormente aplicar el test de Levene.

El contraste de Levene se define como:

$$H_0: \sigma_1 = \sigma_2 = \dots = \sigma_k$$

$$H_1: \sigma_i \neq \sigma_j$$

Dada una variable  $Y$  con tamaño muestral  $N$  dividido en  $k$  individuos o grupos, donde  $N_i$  es el tamaño muestral del individuo o grupo  $i$ , el test de Levene se define como:

$$W = \frac{(N-k) \sum_{i=1}^k N_i (\bar{Z}_{i\cdot} - \bar{Z}_{\cdot\cdot})^2}{(k-1) \sum_{i=1}^k \sum_{j=1}^{N_i} (Z_{ij} - \bar{Z}_{i\cdot})^2}$$

Donde  $Z_{ij} = |Y_{ij} - \bar{Y}_{i\cdot}|$ , siendo  $\bar{Y}_{i\cdot}$  la media del individuo o grupo  $i$ .

Para un nivel de significación  $\alpha$  rechazaremos la hipótesis nula de igualdad de varianzas cuando  $W > F_{(k-1, N-k)}$

#### 4.1.1.1.3.- CORRECCIÓN DE LA HETEROSCEDASTICIDAD.

Una vez detectada la presencia de heteroscedasticidad y partiendo del modelo  $y_{it} = x_{it}\beta + v_{it}$   $v_{it} = c_i + u_{it}$   $t = 1, 2, \dots, T$   $i = 1, 2, \dots, N$  Baltagi, (2001:78) propone la transformación que veremos a continuación para que modelo cumpla con la hipótesis básica de homoscedasticidad.

Dicha transformación es  $y_{it}^* = (y_{it} - \theta_i \bar{y}_{i\cdot})$  donde  $\theta_i = 1 - \left( \frac{\sigma_u}{\tau_i} \right)$  para  $i = 1, 2, \dots, N$  y  $\tau_i^2 = Tw_i^2 + \sigma_u^2$ .

Como puede apreciarse esta transformación propone una modificación personalizada del modelo para cada individuo, de tal forma que dos individuos del mismo panel sufrirán distintas modificaciones. El elemento que particulariza la transformación es  $\tau_i$ .

La aplicación práctica de esta transformación, en otras palabras los MCG Factibles, requieren de la estimación de  $\sigma_u^2$  y  $w_i^2$  para  $i = 1, 2, \dots, N$ . Queda claro inmediatamente que este procedimiento requiere una  $T$  grande y una  $N$  preferiblemente pequeña con  $T > N$ . Esta no es la típica situación de datos de panel pero si el caso de nuestro

trabajo donde nos encontramos sólo con 4 individuos y más de 100 observaciones en el tiempo de cada uno de ellos.

Para el modelo de EFECTOS FIJOS (EF) deberíamos estimar  $\hat{\sigma}_u^2 = \frac{\hat{u}'_i \hat{u}_i}{N(T-1) - K}$

donde  $\hat{u}_i$  son los residuos de la regresión

$$y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)\beta + (u_{it} - \bar{u}_i)$$

$$y_{it} = x_{it}\beta + c_i + u_{it}$$

$$\bar{y}_i = \bar{x}_i\beta + c_i + \bar{u}_i$$

$$\text{donde } \bar{y}_i = \frac{\sum_{t=1}^T y_{it}}{T}, \quad \bar{x}_i = \frac{\sum_{t=1}^T x_{it}}{T} \quad \text{y} \quad \bar{u}_i = \frac{\sum_{t=1}^T u_{it}}{T}.$$

Sabemos que  $\sigma_{v_i}^2 = w_i^2 + \sigma_u^2 \quad i = 1, 2, \dots, N$ . Para estimar  $\sigma_{v_i}^2$  hacemos la regresión de MCO Agrupados sobre el modelo  $y_{it} = x_{it}\beta + v_{it}$  y obtenemos los residuos  $\hat{v}_{it}$  y a

continuación estimamos  $\hat{\sigma}_{v_i}^2 = \frac{\sum_{t=1}^T (\hat{v}_{it} - \hat{v}_{i\cdot})^2}{T-1}$  para  $i = 1, 2, \dots, N$ . Esta estimación toma un valor distinto para cada individuo del panel.

Por último nos queda la estimación de  $w_i$ , sabemos que  $\hat{w}_i^2 = \hat{\sigma}_{v_i}^2 - \hat{\sigma}_u^2 \quad i = 1, 2, \dots, N$ , que de nuevo tomará un valor distinto para cada individuo del panel. Por fin estamos en situación de sustituir todas las estimaciones en la expresión de transformación que elimina la heteroscedasticidad.

En el caso del modelo de EFECTOS ALEATORIOS (EA) la única diferencia se produce en la estimación de  $\hat{\sigma}_u^2 = \frac{\hat{u}'_i \hat{u}_i}{N(T-1) - K}$  donde en este caso los residuos  $\hat{u}_i$

son los residuos procedentes de la estimación del modelo  $y_{it} - \lambda \bar{y}_i = (x_{it} - \lambda \bar{x}_i) \beta + (u_{it} - \lambda \bar{u}_i)$ .

La estimación de  $\sigma_{v_i}^2$  sigue los mismos pasos ya que está basada en el regresión por MCO Agrupados y esta estimación no cambia de un método al otro.

En conclusión esta transformación sigue el mismo esquema que la transformación propuesta en el modelo de efectos fijos, obteniendo valores distintos para cada individuo que lógicamente serán distintos a los valores obtenidos por dichos individuos en la transformación del modelo de efectos fijos.

#### **4.1.2.- AUTOCORRELACIÓN.**

La estimación en presencia de autocorrelación en el término de la perturbación aleatoria produce estimaciones de los parámetros consistentes pero ineficientes y los errores estándar de dichas estimaciones estarán sesgados.

Por tanto, el problema es de tal índole que no queda más remedio que corregir el modelo; ya que en caso contrario el trabajo realizado hasta ese momento no sería válido.

##### **4.1.2.1.- DETECCIÓN DE LA AUTOCORRELACIÓN.**

En este apartado nos acercamos al problema de testar conjuntamente autocorrelación y efectos individuales. Baltagi y Li (1995) derivaron tres estadísticos LM para un modelo de componentes del error, con errores con autocorrelación de orden 1. El primer estadístico LM contrasta conjuntamente autocorrelación de orden cero y efectos individuales aleatorios. El segundo estadístico LM contrasta autocorrelación de orden cero asumiendo efectos individuales fijos y el tercer estadístico testa autocorrelación de orden cero asumiendo efectos individuales aleatorios. En todos estos casos Baltagi y Li (1995) demostraron que el estadístico LM es el mismo tanto si la hipótesis alternativa es un AR(1) o un MA(1). Además Baltagi y Li (1995) derivaron dos extensiones del test de Burke, Godfrey y Termayne (1990) en el que se plantea contrastar si el proceso sigue un AR(1) frente a un MA(1). En el primero de ellos se contrasta la hipótesis de que la perturbación aleatoria sigue un AR(1) frente a un MA(1). En el segundo la hipótesis nula es que la perturbación sigue un MA(1) frente a un AR(1) en un modelo de componentes del error.

Estos contrastes son simples en el sentido de que sólo requieren la obtención de residuos MCO o WITHIN.

Consideraremos la regresión de datos de panel dada por

$$y = \alpha i_{NT} + X\beta + u = Z\delta + u$$

$$y_{it} = z'_{it}\delta + u_{it} \quad i = 1 \dots N \quad t = 1 \dots T$$

donde  $K$  es un vector de  $(K+1) \times 1$  coeficientes de la regresión incluido el término independiente. La perturbación aleatoria sigue “ONE WAY ERROR COMPONENT MODEL”, esto es las diferencias entre las perturbaciones de los individuos se producen precisamente entre los individuos y no porque exista algún componente temporal que haga surgir esas diferencias entre individuos.

$$u_{it} = \mu_i + v_{it}$$

En el que  $\mu_i \sim IIN(0, \sigma_\mu^2)$  y la perturbación restante sigue un proceso AR(1) estacionario:  $v_{it} = \rho v_{it-1} + \varepsilon_{it}$  con  $|\rho| < 1$  o un proceso MA(1):  $v_{it} = \varepsilon_{it} + \lambda \varepsilon_{it-1}$  con  $|\lambda| < 1$  y  $\varepsilon_{it} \sim IIN(0, \sigma_\varepsilon^2)$

#### 4.1.2.1.1.- TEST LM CONJUNTO PARA AUTOCORRELACIÓN Y EFECTOS INDIVIDUALES ALEATORIOS.

Siguiendo a Baltagi (2001:90) vamos a considerar el test LM conjunto para un modelo de componentes del error donde la distribución sigue un proceso MA(1). En este caso la matriz de varianzas y covarianzas de la perturbación es

$$\Omega = E[uu'] = \sigma_\mu^2 I_N \otimes J_T + \sigma_\varepsilon^2 I_N \otimes V_\lambda \quad \text{donde } V_\lambda = \begin{bmatrix} 1+\lambda^2 & \lambda & 0 & 0 \\ \lambda & 1+\lambda^2 & \lambda & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1+\lambda^2 \end{bmatrix} \quad \text{y el logaritmo}$$

de la función de verosimilitud viene dado por  $L(\delta, \theta) = \text{constante} - \frac{1}{2} \log |\Omega| - \frac{1}{2} u' \Omega^{-1} u$

con  $\theta = (\lambda, \sigma_\mu^2, \sigma_\varepsilon^2)'$

El test estadístico LM preciso para contrastar  $H_1^a : \sigma_\mu^2 = 0; \quad \lambda = 0$  es:



$$LM_1 = \frac{NT^2}{2(T-1)(T-2)} [A^2 - 4AB + 2TB^2] \quad \text{en el que} \quad A = \left[ \frac{\hat{u}'(I_N \otimes J_T)\hat{u}}{\hat{u}'\hat{u}} \right]^{-1} \quad y$$

$$B = \frac{\sum_{i=1}^N \sum_{t=2}^T \hat{u}_{it} \hat{u}_{it-1}}{\hat{u}'\hat{u}} \quad \text{siendo } \hat{u} \text{ los residuos MCO.}$$

Este estadístico se distribuye asintóticamente (para N grande) como una  $\chi^2_2$  bajo  $H_1^a$ . Baltagi y Li (1991b) demuestra que este contraste estadístico es exactamente el mismo que el contraste necesario para testar la hipótesis:  $H_1^b : \sigma_\mu^2 = 0; \rho = 0$  en el que la perturbación aleatoria sigue un proceso AR(1). Por lo que el test estadístico conjunto LM para efectos individuales aleatorios y autocorrelación de primer orden es independiente de la forma de la autocorrelación, tanto si es un AR(1) o un MA(1).

Asumiendo que no existe autocorrelación podemos adaptar este test para contrastar la hipótesis  $H_2 : \sigma_\mu^2 = 0$ . En este caso el estadístico será  $LM_2 = \sqrt{\frac{NT}{2(T-1)}} \cdot A$  que está asintóticamente distribuido (para N grande) como una  $N(0,1)$  bajo  $H_2$  contra la alternativa unilateral  $H_2' : \sigma_\mu^2 > 0$ .

Una nueva transformación de este estadístico nos permite contrastar la hipótesis  $H_3 : \rho = 0$  (o en su caso  $\lambda = 0$ ) asumiendo que no existen efectos individuales. En este caso el estadístico será  $LM_3 = \sqrt{\frac{NT^2}{T-1}} \cdot B$  que se distribuye asintóticamente (para N grande) como  $N(0,1)$  bajo  $H_3$  contra la hipótesis alternativa unilateral  $H_3' : \rho > 0$  (o en su caso  $\lambda > 0$ ).

La presencia del término de interacción (el producto AB) en el test estadístico conjunto LM enfatiza la importancia de este test cuando se sospecha la presencia de autocorrelación y efectos individuales aleatorios. Sin embargo cuando T es grande el término interacción se vuelve insignificante.

En todos estos test LM se ha asumido que la hipótesis nula subyacente es que la perturbación aleatoria sigue un proceso ruido blanco. Sin embargo, en aplicaciones empíricas de datos de panel es habitual considerar la presencia de efectos individuales. En estos casos no resulta adecuado contrastar la autocorrelación asumiendo la NO existencia de efectos individuales como se hace en  $H_3$ . Si utilizamos el estadístico  $LM_3$  para testar autocorrelación se está muy cerca de rechazar la hipótesis nula  $H_3$  incluso si la hipótesis es cierta. Esto sucede porque  $\mu_i$  está correlacionado para el mismo individuo a lo largo del tiempo y contribuirá a rechazar la hipótesis nula de NO autocorrelación.

#### 4.1.2.1.2.- TEST LM CONJUNTO PARA AUTOCORRELACIÓN DE PRIMER ORDEN EN UN MODELO DE EFECTOS ALEATORIOS.

Baltagi y Li (1995) derivaron un test LM para autocorrelación de primer orden, suponiendo que existan efectos individuales aleatorios. En el caso de un modelo AR(1), la hipótesis nula es:

$$H_4^b: \rho = 0 \quad (\text{dado } \sigma_\mu^2 > 0)$$

$$H_4^{b'}: \rho \neq 0 \quad (\text{dado } \sigma_\mu^2 > 0)$$

La matriz de varianzas y covarianzas (bajo la hipótesis alternativa) es  $\Omega_1 = \sigma_\mu^2 (I_N \otimes J_T) + \sigma_\varepsilon^2 (I_N \otimes V_\rho)$ , bajo la hipótesis nula  $H_4^b$ , tenemos

$$(\Omega_1^{-1}) /_{\rho=0} = \left( \frac{1}{\sigma_\varepsilon^2} \right) I_N \otimes E_T + \left( \frac{1}{\sigma_1^2} \right) I_N \otimes \bar{J}_T$$

$$\left( \frac{\partial \Omega_1}{\partial \rho} \right) /_{\rho=0} = \sigma_\varepsilon^2 (I_N \otimes G)$$

$$\left( \frac{\partial \Omega_1}{\partial \sigma_\mu^2} \right) /_{\rho=0} = (I_N \otimes J_T)$$

$$\left( \frac{\partial \Omega_1}{\partial \sigma_\varepsilon^2} \right) \Big|_{\rho=0} = (I_N \otimes I_T)$$

Donde  $\bar{J}_T = \iota_T \iota_T' / T$ ,  $\iota_T$  es un vector de unos de dimensión  $T$ ;  $E_T = I_T - \bar{J}_T$ ,  $G$  es una matriz bidiagonal con los elementos de la bidiagonal todos igual a 1 y  $\sigma_1^2 = T\sigma_\mu^2 + \sigma_\varepsilon^2$ ; utilizando estos resultados obtenemos

$$D_\rho = \left( \frac{\partial L}{\partial \rho} \right) \Big|_{\rho=0} = \left( \frac{N(T-1)}{T} \right) \bullet \frac{\sigma_1^2 - \sigma_\varepsilon^2}{\sigma_1^2} + \left( \frac{\sigma_\varepsilon^2}{2} \right) u' \left[ I_N \otimes \left( \frac{\bar{J}_T}{\sigma_1^2} + \frac{E_T}{\sigma_\varepsilon^2} \right) G \left( \frac{\bar{J}_T}{\sigma_1^2} + \frac{E_T}{\sigma_\varepsilon^2} \right) \right] u$$

La matriz de información es

$$\hat{J} = \begin{bmatrix} \hat{J}_{\rho\rho} & \frac{N(T-1)\hat{\sigma}_\varepsilon^2}{\hat{\sigma}_1^4} & \frac{N(T-1)}{T}\hat{\sigma}_\varepsilon^2 \left[ \frac{1}{\hat{\sigma}_1^4} - \frac{1}{\hat{\sigma}_\varepsilon^4} \right] \\ \left( \frac{NT^2}{2\hat{\sigma}_1^4} \right) & \frac{NT}{2\hat{\sigma}_1^4} & \\ \frac{N}{2} \left[ \frac{1}{\hat{\sigma}_1^4} + \frac{T-1}{\hat{\sigma}_\varepsilon^4} \right] & & \end{bmatrix}$$

Donde  $\hat{J}_{\rho\rho} = N \left[ 2a^2(T-1)^2 + 2a(2T-3) + (T-1) \right]$

$$a = \frac{\hat{\sigma}_\varepsilon^2 - \hat{\sigma}_1^2}{T\hat{\sigma}_1^2}$$

Por tanto el test estadístico LM es  $LM = \hat{D}'\hat{J}^{-1}\hat{D} = \left( \hat{D}_\rho \right)^2 \hat{J}''$  donde  $\hat{J}'' = \frac{N^2 T^2 (T-1)}{\det(\hat{J}) 4\hat{\sigma}_1^4 \hat{\sigma}_\varepsilon^4}$ ,

bajo la hipótesis nula  $H_4^b$ , el contraste LM se distribuye asintóticamente (para  $N$  grande) como una  $\chi_1^2$ .

$$\hat{\sigma}_\varepsilon^2 = \frac{\hat{u}'(I_N \otimes E_T)\hat{u}}{N(T-1)}$$

$$\hat{\sigma}_1^2 = \frac{\hat{u}'(I_N \otimes \bar{J}_T)\hat{u}}{N}$$

Donde  $\hat{u}$  son los residuos máximo verosímiles bajo la hipótesis nula.

El test LM unilateral para la hipótesis  $H_4^b$  (correspondiente a la alternativa  $\rho > 0$ ) es

$LM_4 = \hat{D}_\rho \sqrt{\hat{J}''}$  y se distribuye asintóticamente (para N grande) como una  $N(0,1)$ .

Cuando la autocorrelación es del tipo MA(1), la hipótesis nula se convierte en  $H_4^a: \lambda = 0$  (dado  $\sigma_\mu^2 > 0$ ) frente a  $H_4^{a'}: \lambda \neq 0$  (dado  $\sigma_\mu^2 > 0$ ). En este caso la matriz de varianzas y covarianzas es  $\Omega_2 = \sigma_\mu^2 (I_N \otimes J_T) + \sigma_\varepsilon^2 (I_N \otimes V_\lambda)$ .

El test estadístico es el mismo que en el caso anterior.

Para resumir, el test LM para testar autocorrelación de primer orden, asumiendo efectos individuales aleatorios, es invariable a la forma de la autocorrelación (si es AR(1) o MA(1)).

Tal y como nos muestra Baltaggi (2001: 93), Bera, Sosa – Escudero y Yoon (2001) sugieren un ajuste al test LM que lo convierte en robusto frente a un problema de especificación local.

En este caso  $LM_\mu = \frac{NTA^2}{2(T-1)}$  para testar  $H_2: \sigma_\mu^2 = 0$  ignorando la posible presencia de autocorrelación.

Los autores sugieren calcular  $LM_\mu^* = \frac{NT(2B-A)^2}{2(T-1)\left(1-\frac{2}{T}\right)}$ . Este test esencialmente

modifica  $LM_\mu$  corrigiendo la media y varianza de  $\frac{\partial L}{\partial \sigma_\mu^2}$  por su correlación asintótica con  $\frac{\partial L}{\partial \rho}$ . Bajo la hipótesis nula,  $LM_\mu^*$  se distribuye asintóticamente como una  $\chi_1^2$ .

De manera similar, los autores sugieren calcular  $LM_{\rho}^* = \frac{NT^2 \left(B - \frac{A}{T}\right)^2}{(T-1) \left(1 - \frac{2}{T}\right)}$  en lugar de

$LM_{\rho} = \frac{NT^2 B^2}{(T-1)}$  para testar  $H_3: \rho=0$  frente a la alternativa que  $\rho \neq 0$  ignorando la presencia de efectos individuales aleatorios.

#### 4.1.2.1.3.- TEST LM CONJUNTO PARA AUTOCORRELACIÓN DE PRIMER ORDEN EN UN MODELO DE EFECTOS FIJOS.

Consideremos el modelo  $y_{it} = z'_{it}\delta + u_{it} \quad i=1 \dots N \quad t=1 \dots T$  en el que  $\delta$  es un vector de  $(k+1) \times 1$  coeficientes incluyendo la constante.

$$u_{it} = \mu_i + v_{it}$$

La hipótesis nula es  $H_5^b: \rho=0$  supuesto que  $\mu_i$  sea un parámetro. Escribiendo las variables de cada individuo como un vector de dimensión  $T \times 1$ , tenemos:

$$y_i = z_i \delta + \mu_i l_T + v_i$$

$$y_i = (y_{i1}, y_{i2}, \dots, y_{iT})$$

donde  $z_i$  es  $T \times (k+1)$  en el que  $\Omega_{\rho} = \sigma_{\varepsilon}^2 v_{\rho}$  siguen un proceso AR(1).

$$v_i \text{ es } T \times 1; \quad v_i \sim N(0, \Omega_{\rho})$$

El logaritmo de la función de verosimilitud es

$$L(\delta, \rho, \mu, \sigma_{\varepsilon}^2) = \text{constante} - \frac{1}{2} \log |\Omega| - \frac{1}{2\sigma_{\varepsilon}^2} \sum_{i=1}^N (y_i - z_i \delta - \mu_i l_T)' V_{\rho}^{-1} (y_i - z_i \delta - \mu_i l_T) \text{ en el que}$$

$\Omega = I_N \otimes \Omega_{\rho}$  es la matriz de varianzas y covarianzas de  $v' = (v'_1, \dots, v'_N)$ .

Se puede comprobar que el estimador máximo verosímil de  $\mu_i$  viene dado por:

$\hat{\mu}_i = \left\{ \left( i_T' V_\rho^{-1} i_T \right)^{-1} \left[ i_T' V_\rho^{-1} \left( y_i - z_i \hat{\delta} \right) \right] \right\}_{\rho=0} = \bar{y}_{i\cdot} - \bar{z}_{i\cdot}' \hat{\delta}$  donde  $\hat{\delta}$  es el estimador máximo

verosímil de  $\delta$ ,  $\bar{y}_{i\cdot} = \frac{\sum_{t=1}^T y_{it}}{T}$  y  $\bar{z}_{i\cdot}$  es un vector de  $(K+1) \times 1$  de promedios de  $z_{it}$  en el tiempo.

Escribimos el logaritmo de la función de verosimilitud en forma vectorial de  $V$  como:

$$L(\delta, \mu, \vartheta) = \text{constante} - \frac{1}{2} \log |\Omega| - \frac{1}{2} v' \Omega^{-1} v \text{ en el que } \vartheta' = (\rho, \sigma_\varepsilon^2)$$

A partir de esta expresión puede observarse que el test LM para contrastar la hipótesis  $H_5^b$  es:

$LM = \left( \frac{NT^2}{T-1} \right) \left( \frac{\hat{v}' \hat{v}_{-1}}{\hat{v}' \hat{v}} \right)^2$  que se distribuye asintóticamente (para una  $T$  grande) como una  $\chi_1^2$  bajo la hipótesis nula  $H_5^b$ .

Nótese que  $\hat{v}_{it} = y_{it} - z_{it}' \hat{\delta} - \hat{\mu}_i = (\tilde{y}_{it} - \tilde{z}_{it}' \hat{\delta}) + (\bar{y}_{i\cdot} - \bar{z}_{i\cdot}' \hat{\delta} - \hat{\mu}_i)$  donde  $\tilde{y}_{it} = y_{it} - \bar{y}_{i\cdot}$  es la transformación habitual Within.

Bajo la hipótesis nula de  $\rho=0$ , el último término del paréntesis es cero dado que  $\{\hat{\mu}\}_{\rho=0} = \bar{y}_{i\cdot} - \bar{z}_{i\cdot}' \hat{\delta}$  y  $\{\hat{v}_{it}\}_{\rho=0} = \tilde{y}_{it} - \tilde{z}_{it}' \hat{\delta} = \tilde{v}_{it}$ . Esto significa que el estadístico LM dado anteriormente puede expresarse en términos de los residuos WITHIN habituales (los  $\tilde{v}$ ).

El test unilateral para contrastar la hipótesis  $H_5^b$  (correspondiente a la alternativa

$\rho > 0$ ) es  $LM_5 = \sqrt{\frac{NT^2}{T-1}} \cdot \frac{\tilde{v}' \tilde{v}_{-1}}{\tilde{v}' \tilde{v}}$  que se distribuye asintóticamente (para una  $T$  grande) como una Normal  $(0,1)$ .

Con argumentos similares, puede demostrarse que el test estadístico para  $H_5^b : \lambda = 0$ , en un modelo de efectos fijos con perturbaciones que sigan un MA(1), es idéntico al  $LM_5$ .

Es de hacer notar que  $LM_5$  difiere de  $LM_3$  exclusivamente en el hecho de que los residuos WITHIN  $\tilde{v}$  (en  $LM_5$ ) reemplazan a los residuos MCO  $\hat{u}$  (en  $LM_3$ ). Dado que la transformación WITHIN elimina los efectos individuales tanto si son fijos o aleatorios, puede utilizarse el contraste  $LM_5$  para contrastar la autocorrelación en los modelos de efectos aleatorios.

#### 4.1.2.1.4.- TEST DURBIN – WATSON PARA DATOS DE PANEL.

Para el modelo de efectos fijos  $y_i = z_i\delta + \mu_i t_T + v_i$  con  $v_i$  siguiendo un proceso AR(1), Bhargava, Franzini y Narendranathan (1982) sugirieron un test  $H_0 : \rho = 0$  frente a la alternativa  $|\rho| < 1$ , utilizando el estadístico de Durbin – Watson basado en los residuos WITHIN ( $\tilde{v}_{it}$ ) en lugar de los residuos MCO.

$$d_\rho = \frac{\sum_{i=1}^N \sum_{t=2}^T (\tilde{v}_{it} - \tilde{v}_{it-1})^2}{\sum_{i=1}^N \sum_{t=1}^T \tilde{v}_{it}^2}$$

Bhargava, Franzini y Narendranathan demostraron que para regresores arbitrarios,  $d_\rho$  es un test más poderoso en la cercanía de  $\rho = 0$ . Los autores argumentan que valores críticos exactos del estadístico son impracticables e innecesarios para datos de panel dado que involucran el cálculo de autovalores no cero para matrices de grandes dimensiones ( $NT \times NT$ ).

Los autores nos muestran cómo pueden calcularse las fronteras superior e inferior de  $d_\rho$  y tabulan al nivel del 5% para  $N=50, 100, 150, 250, 500, 1000$ ,  $T=6, 10$ , y  $K=1, 3, 5, 7, 9, 11, 13, 15$ .

Bhargava, Franzini y Narendranathan remarcan que  $d_\rho$  será raramente no concluyente dado que las fronteras serán muy delgadas incluso para valores moderados de  $N$ . También para  $N$  grandes, es más, argumentan que no es necesario calcular esta frontera, basta simplemente con ver si  $d_\rho$  es menor que 2 cuando se está testando autocorrelación positiva.

Los mismos autores también nos sugieren utilizar el estadístico Berenblut – Webb para testar  $H_0: \rho=0$  porque es un test invariante localmente más poderoso en la cercanía de  $\rho=1$ . Este estadístico viene dado por:

$$g_\rho = \frac{\sum_{i=1}^N \sum_{t=2}^T \Delta \hat{u}_{it}^2}{\sum_{i=1}^N \sum_{t=1}^T \tilde{v}_{it}^2} \text{ donde } \Delta \hat{u}_{it} \text{ son los residuos MCO obtenidos de la versión de primeras}$$

diferencias de la ecuación de regresión

$$y_{it} = z'_{it} + u_{it}$$

$$u_{it} = \mu_i + v_{it}$$

y  $\tilde{v}_{it}$  son los residuos WITHIN. BFN demuestran que  $g_\rho$  y  $d_\rho$  tienen un poder similar cuando  $N=30, T=10, K=2, \alpha=0,05$  y  $\rho=0,25; 0,4; 0,5$ . También que ambos test son equivalentes si  $N$  es grande.

Bhargava, Franzini y Narendranathan también proponen un test para contrastar si los residuos siguen un Paseo Aleatorio; esto es,  $H_0: \rho=1$  frente a la alternativa  $|\rho| < 1$ . Esta prueba está basada en el estadístico  $R_\rho = \frac{\Delta \hat{u}' \Delta \hat{u}}{(\Delta \hat{u})' F^* (\Delta \hat{u})}$  donde

$\Delta \hat{u}$  son los residuos MCO utilizados en el cálculo del estadístico  $g_\rho$ ,  $F^* = I_N \otimes F$



siendo  $F$  una matriz simétrica de dimensión  $(T-1) \times (T-1)$  con los elementos dados por  $F_{j,s} = \frac{(T-j)s}{T}$  si  $j \geq s$  ( $j, s = 1, \dots, T-1$ ).

Para regresores generales, BFN demuestran que  $R_\rho \leq g_\rho \leq d_\rho$  donde  $g_\rho$  y  $d_\rho$  están siendo considerados bajo la hipótesis nula de Paseo Aleatorio. Los autores tabulan al 5 % las fronteras superior e inferior para  $R_\rho$  y sugieren que las fronteras para  $R_\rho$  puedan utilizarse en la práctica para  $g_\rho$  y  $d_\rho$ . Sin embargo cuando  $N \rightarrow \infty$  como es típico en paneles, los tres contrastes son equivalentes;  $R_\rho = g_\rho = d_\rho$  y BFN recomiendan en ese caso sólo el test de Durbin – Watson  $d_\rho$  calculado para testar la hipótesis de Paseo Aleatorio.

#### **4.1.2.1.5.- TEST PARA DIFERENCIAR UN AR(1) FRENTE UN MA(1) EN UN MODELO DE COMPONENTES DEL ERROR.**

Burkey, Godfrey y Termayne (1990) derivaron un test simple para distinguir entre un proceso AR(1) y un MA(1). Baltagi y Li (1995) propusieron dos extensiones del test de BGT en el modelo de componentes del error. Estos contrastes son simples de implementar y requieren para su cálculo residuos WITHIN y MCO.

La idea básica del test de BGT es: bajo la hipótesis nula de un proceso AR(1), el error  $v_{it}$  satisface la condición:  $correl(v_{it}, v_{i,t-\tau}) = \rho_\tau = (\rho_1)^\tau$   $\tau = 1, 2, \dots$  por lo tanto bajo la hipótesis nula, se tiene que cumplir que  $\rho_2 - (\rho_1)^2 = 0$ . Bajo la hipótesis alternativa de un proceso MA(1) sobre  $v_{it}$ ,  $\rho_2 = 0$  y por tanto  $\rho_2 - (\rho_1)^2 \leq 0$ .

Burkey, Godfrey y Termayne recomiendan un test basado en  $\rho_2 - (\rho_1)^2 = 0$  usando estimaciones de  $\rho$  obtenidas de residuos MCO. Sin embargo un problema

permanece; King (1983) sugirió que cualquier “buen” test debería tener un tamaño que tendiera a 0 asintóticamente, para  $\rho > 0,5$ .

Este contraste no garantiza el cumplimiento de esta propiedad. Para remediarlo, los autores propusieron implementar es test anterior con la decisión de aceptar la hipótesis nula de AR(1) si  $\hat{\rho}_1 > \frac{1}{2} + \frac{1}{\sqrt{T}}$ .

En un modelo de componentes del error, la transformación WITHIN elimina los efectos individuales y se pueden utilizar los residuos WITHIN de  $\tilde{u}_{i,t} (= \tilde{v}_{i,t})$  en lugar de los residuos MCO para construir el test BGT.

$$\text{Sea } (\tilde{\rho}_1)_i = \frac{\sum_{t=2}^T \tilde{u}_{i,t} \tilde{u}_{i,t-1}}{\sum_{t=1}^T \tilde{u}_{i,t}^2} \text{ y } (\tilde{\rho}_2)_i = \frac{\sum_{t=3}^T \tilde{u}_{i,t} \tilde{u}_{i,t-2}}{\sum_{t=1}^T \tilde{u}_{i,t}^2} \quad i = 1, \dots, N$$

El siguiente test, basado en el de BGT y que encontramos en Baltagi (2001:97) es:

$$\tilde{\gamma}_i = \frac{\sqrt{T} [(\tilde{\rho}_2)_i - (\tilde{\rho}_1^2)_i]}{1 - (\tilde{\rho}_2)_i} \text{ se distribuye asintóticamente (para } T \text{ grande) como una}$$

normal (0,1) bajo la hipótesis nula de un AR(1). Utilizando los datos de los N individuos, podemos construir un test BGT generalizado para el modelo de componentes del error.

$$\tilde{\gamma} = \sqrt{N} \left( \frac{\sum_{i=1}^N \tilde{\gamma}_i}{N} \right) = \sqrt{NT} \frac{\sum_{i=1}^N \left[ \frac{(\tilde{\rho}_2)_i - (\tilde{\rho}_1^2)_i}{1 - (\tilde{\rho}_2)_i} \right]}{N}$$

En esta expresión  $\tilde{\gamma}_i$  es independiente de los diferentes individuos ( $i$ ) ya que los residuos  $\tilde{u}_i$  son independientes. De ahí que  $\tilde{\gamma}$  se distribuya asintóticamente (para  $t$  grande) como una Normal (0,1) bajo la hipótesis nula de un proceso AR(1).

Este último contraste se complementa con el siguiente test:

$$\tilde{r}_1 = \frac{\sum_{i=1}^N (\tilde{r}_1)_i}{N} = \frac{1}{N} \sum_{i=1}^N \left[ \frac{\sum_{t=2}^T \tilde{u}_{i,t} \tilde{u}_{i,t-1}}{\sum_{t=1}^T \tilde{u}_{i,t}^2} \right]$$

En conclusión Baltagi y Li (1995) propusieron un contraste BGT que puede sintetizarse de la siguiente forma:

- 1) Utilizar los residuos WITHIN  $\tilde{u}_{i,t}$  para calcular  $\tilde{\gamma}$  y  $\tilde{r}_1$ .
- 2) Aceptar el modelo AR(1) si  $\tilde{\gamma} > c_\alpha$  o  $\tilde{r}_1 > \frac{1}{2} + \frac{1}{\sqrt{T}}$  donde  $\Pr[N(0,1) \leq c_\alpha] = \alpha$

El sesgo al estimar  $\rho_s (s=1,2)$  utilizando residuos WITHIN es de  $O\left(\frac{1}{T}\right)$  cuando  $N \rightarrow \infty$  (Nickell, 1981) Por lo tanto BGT<sub>1</sub> puede no funcionar bien para  $T$  pequeñas. Dado que en el típico panel  $N$  es grande y  $T$  pequeña, sería deseable un test alternativo simple que funcionara bien con  $N$  grande en lugar de  $T$  grande. Sin embargo, resulta un contraste adecuado para casos como el que se expone en esta tesis, en la que  $T$  es grande y  $N$  pequeña

#### 4.1.2.1.6.- TEST ALTERNATIVO TIPO BGT PARA CONTRASTAR UN AR(1) FRENTE UN MA(1).

Sea la hipótesis nula  $H_\tau : v_{i,t} = \varepsilon_{i,t} + \lambda \varepsilon_{i,t-1}$  y la alternativa  $H'_\tau : v_{i,t} = \rho v_{i,t} + \varepsilon_{i,t}$  donde  $\varepsilon_{i,t} \sim N(0, \sigma_\varepsilon^2)$ . Nótese que este test difiere del anterior en que la hipótesis nula es un MA(1) en lugar de un AR(1).

Este test alternativo, tipo BGT, utiliza estimaciones de autocorrelación derivadas de residuos MCO.

Sea:

$$Q_0 = \frac{\sum \sum u_{i,t}^2}{NT} = \frac{u'u}{NT}$$

y

$$Q_s = \frac{\sum \sum u_{i,t} u_{i,t-s}}{N(T-s)} = \frac{u'(I_N \otimes G_s)u}{N(T-s)} \text{ para } s=1, \dots, S$$

Donde  $G_s = \frac{1}{2}$  matriz de Toeplitz( $\iota_s$ ),  $\iota_s$  es un vector de ceros con los  $(s+1)$  elementos siendo uno.  $s=1, \dots, S$  con  $S \leq (T-1)$  y  $S$  finita.

Dados los verdaderos residuos ( $u$ ) y asumiendo que:

$$\left[ \frac{u'Au}{n} - E\left(\frac{u'Au}{n}\right) \right] \xrightarrow{P} 0$$

Donde  $n=NT$  y  $A$  una matriz simétrica arbitraria, Baltagi y Li (1995) probaron los resultados siguientes cuando  $N \rightarrow \infty$

(1) Para el modelo MA(1)

$$\begin{aligned} p \lim Q_0 &= \sigma_u^2 + \sigma_v^2 = \sigma_\varepsilon^2 (1 + \lambda^2) \\ p \lim Q_1 &= \sigma_u^2 + \lambda \sigma_\varepsilon^2 \\ p \lim Q_s &= \sigma_u^2 \quad \text{para } s=2, \dots, S \end{aligned}$$

(2) Para el modelo AR(1)

$$p \lim Q_0 = \sigma_u^2 + \sigma_v^2$$

$$p \lim Q_s = \sigma_u^2 + \rho^s \sigma_v^2 \quad \text{para } s = 1, \dots, S$$

Baltagi y Li (1995) demostraron que para  $N$  grande es posible distinguir entre el proceso AR(1) y el MA(1) basándose en la información obtenida del hecho de que  $Q_s - Q_{s+l}$  para  $s \geq 2$  y  $l \geq 1$ . Para verlo, nótese que  $p \lim(Q_s - Q_{s+l}) = 0$  para el proceso MA(1) y  $p \lim(Q_s - Q_{s+l}) = \sigma_v^2 \rho^s (1 - \rho^l) > 0$  para el proceso AR(1).

A partir de aquí, Baltagi y Li (1995) sugieren un test asintótico de  $H_\tau$  frente a  $H'_\tau$  basado en

$$\gamma = \sqrt{N/V} (Q_2 - Q_3)$$

donde bajo ciertas condiciones de regularidad,  $\gamma$  se distribuye asintóticamente (para  $N$  grande) como una  $N(0,1)$  bajo la hipótesis nula de un proceso MA(1). Para poder calcular  $V$  nótese que para el proceso MA(1),  $\sigma_v^2 = \sigma_\varepsilon^2 (1 + \lambda^2)$  y  $\sigma_\varepsilon^2 V_\lambda = \sigma_v^2 I_T + \sigma_\varepsilon^2 \lambda G$ . Por tanto, no necesitamos estimar  $\lambda$  para calcular el test  $\gamma$ ; todo lo que necesitamos es obtener unos estimadores consistentes de  $\sigma_v^2$ ,  $\lambda \sigma_\varepsilon^2$  y  $\sigma_\mu^2$ . Estos estimadores son:

$$\hat{\sigma}_v^2 = \hat{Q}_0 - \hat{Q}_2$$

$$\lambda \hat{\sigma}_\varepsilon^2 = \hat{Q}_0 - \hat{Q}_1$$

$$\hat{\sigma}_\mu^2 = \hat{Q}_2$$

donde  $\hat{Q}_s$  se obtiene de  $Q_s$  reemplazando  $u_{i,t}$  por los residuos MCO  $\hat{u}_{i,t}$ . Sustituyendo estos estimadores consistentes en  $V$  obtenemos  $\hat{V}$  y el test estadístico  $\gamma$  se convierte en

$$\hat{\gamma} = \sqrt{N/\hat{V}} (\hat{Q}_2 - \hat{Q}_3)$$

en el que

$$(\hat{Q}_2 - \hat{Q}_3) = \frac{\sum_{i=1}^N \sum_{t=3}^T \hat{u}_{i,t} \hat{u}_{i,t-2}}{N(T-2)} - \frac{\sum_{i=1}^N \sum_{t=4}^T \hat{u}_{i,t} \hat{u}_{i,t-3}}{N(T-3)}$$

y

$$\hat{V} = 2tr \left\{ \left[ \left( \hat{\sigma}_\mu^2 J_T + \hat{\sigma}_v^2 I_T + \sigma_\varepsilon^2 \hat{\lambda} G \right) \left( \frac{G_2}{(T-2)} + \frac{G_3}{(T-3)} \right) \right]^2 \right\}$$

$\hat{\gamma}$  se distribuye asintóticamente (para  $N$  grande) como una  $N(0,1)$  bajo la hipótesis nula  $H_\tau$  y conocido como el test BGT<sub>2</sub>.

Baltagi y Li (1995) realizaron experimentos de Monte Carlo, sin embargo la distribución resultante no permite decir si sigue un proceso AR(1) o MA(1).

Los principales resultados pueden resumirse del siguiente modo:

- 1) El test  $LM_1$  conjunto funciona bien al contrastar la hipótesis nula  $H_1 : \rho = \sigma_\mu^2 = 0$ . Su tamaño estimado no es estadísticamente diferente de su tamaño nominal. Sea  $w = \sigma_u^2 / \sigma^2$  la proporción de la varianza total debida al efecto individual. Baltagi y Li (1995) encontraron que en presencia de efectos individuales grandes ( $w > 0,2$ ) o elevada autocorrelación  $\rho$  (o  $\lambda$ )  $> 0,2$ , el test  $LM_1$  tiene un elevado poder, rechazando la hipótesis nula en el 99-100 % de los casos. Su poder es bajo cuando  $w=0$  y  $\rho$  (o  $\lambda$ )  $> 0,2$  o cuando  $w=0,2$  y  $\rho$  (o  $\lambda$ )  $= 0$ .
- 2) El test estadístico  $LM_2$  para contrastar la hipótesis nula  $H_2 : \sigma_\mu^2 = 0$  implícitamente asume que  $\rho$  (o  $\lambda$ )  $= 0$ . Cuando  $\rho$  es de verdad igual a 0, este test funciona bien. Sin embargo, como  $\rho$  se mueve desde 0 y puede crecer hasta 1, este test tiende a estar sesgado a favor de rechazar la hipótesis nula. Esto es porque un coeficiente de autocorrelación elevado ( $\rho$  grande) contribuye a una correlación alta entre los individuos de la muestra,

incluso cuando  $\sigma_\mu^2 = 0$ . Por ejemplo, cuando la hipótesis nula es cierta ( $\sigma_\mu^2 = 0$ ) para  $\rho = 0,9$ ,  $LM_2$  la rechaza en el 100 % de los casos. Resultados similares se obtienen en el caso de que  $v_{i,t}$  siga un proceso MA(1). En general, la presencia de autocorrelación positiva tiende a sesgar el caso a favor de encontrar que los efectos individuales no son cero.

- 3) De forma similar, el test  $LM_3$  para testar  $H_3: \rho = 0$  implícitamente asume que  $\sigma_\mu^2 = 0$ . Este test funciona bien cuando efectivamente  $\sigma_\mu^2 = 0$ . Sin embargo cuando  $\sigma_\mu^2$  aumenta, el funcionamiento del test empeora. Por ejemplo, cuando la hipótesis nula es cierta ( $\rho = 0$ ) pero  $w = 0,9$ ,  $LM_3$  rechaza la hipótesis nula en el 100 % de los casos. Correlación alta entre las  $\mu_i$  contribuye a rechazar la hipótesis nula de no autocorrelación. Estos resultados indican que no se debería ignorar los efectos individuales cuando se contrasta la autocorrelación.
- 4) En contraposición al test  $LM_3$ , los contrastes  $LM_4$  y  $LM_5$  tienen en cuenta los efectos individuales. Para valores elevados de  $\rho$  o  $\lambda$  (mayores que 0,4) tanto  $LM_4$  como  $LM_5$  tienen una potencia de contraste elevada, rechazando la hipótesis nula en más del 99 % de las veces. Sin embargo, el tamaño estimado de  $LM_4$  está más cerca del 5 % del valor nominal que el de  $LM_5$ . Además Baltagi y Li (1995) demuestran que el test modificado de Durbin – Watson de Bhargava, Franzini y Narendranathan (1982) funciona mejor que el  $LM_5$  y lo recomiendan.
- 5) El test Bhargava, Franzini y Narendranathan que utiliza residuos WITHIN y contrasta la hipótesis nula de un AR(1) frente a un MA(1), funciona bien si  $T \geq 60$  y  $T \geq N$ . Sin embargo, cuando  $T$  es pequeño o  $T$  tiene un tamaño moderado respecto a una  $N$  grande,  $BGT_1$  tenderá a sobre rechazar la hipótesis nula. Por tanto  $BGT_1$  no es recomendable en estos casos. Cuando  $N$  es grande y  $T$  pequeño, Baltagi y Li recomiendan el test  $BGT_2$  que utiliza

residuos MCO y contrasta la hipótesis nula de un MA(1) frente a un AR(1). Este test funciona bien cuando  $N$  es grande y no hay confianza en que  $T$  alcance su distribución asintótica. Los resultados Monte Carlo muestran que el funcionamiento de  $BGT_2$  mejora cuando tanto  $N$  o  $T$  aumentan.

Baltagi y Li (1995) realizaron experimentos Monte Carlo para comparar la eficiencia relativa en una muestra finita de un número de estimadores puros y pretestados para un modelo de componentes de error cuya perturbación había sido generada por un proceso AR(1) o MA(1). Estos estimadores eran:

- 1) MCO Mínimos Cuadrados Ordinarios.
- 2) El estimador WITHIN
- 3) El estimador convencional MCG que ignora la autocorrelación de los residuos pero tiene en cuenta la estructura aleatoria de los componentes del error.
- 4) Estimación MCG asumiendo componentes del error aleatorio con la distribución del error siguiendo un proceso MA(1). Lo denotaremos MCGM.
- 5) Estimación MCG asumiendo componentes del error aleatorios con la distribución del error siguiendo un proceso AR(1). Lo denotaremos MCGA.
- 6) Un estimador pretestado basado en los resultados de dos test. Lo denotaremos PRE. El primer test es  $LM_4$  que contrasta la presencia de autocorrelación dada la existencia de efectos individuales aleatorios. Si la hipótesis nula no es rechazada, este estimador se convierte en el tradicional MCG. En el caso de encontrar autocorrelación, se realiza el test  $BGT_2$  para distinguir entre un proceso AR(1) o MA(1) y MCGA o MCGM será el que funcione.
- 7) Un estimador generalizado por el método de los momentos GMM, en el que la estructura de la distribución de los componentes del error es ignorada y la matriz general de varianzas y covarianzas se estima en la dimensión temporal.
- 8) El verdadero estimador MCG, denotado por VMCG, se obtiene por comparación entre los propuestos. En efecto la eficiencia relativa de cada estimador se obtiene dividiendo su ECM (error cuadrático medio) entre VMCG (Verdadero MCG). Es importante enfatizar que todos los estimadores



considerados son consistentes mientras que las variables explicativas y las perturbaciones no estén correlacionadas, cuando  $N \rightarrow \infty$  con  $T$  fijo. La preocupación primaria son sus propiedades en muestras pequeñas. Los resultados muestran que el procedimiento MCG correcto es siempre el mejor, pero los investigadores no tienen un conocimiento perfecto sobre cuál de ellos es: MCGA para un proceso AR(1) o MCGM para un proceso MA(1). En este caso, el estimador pretestado es una alternativa viable dado que si funciona el segundo, está más cerca del correcto MCG que si el verdadero proceso de autocorrelación es un AR(1) o un MA(1).

#### 4.1.2.2.- CORRECCIÓN DE LA AUTOCORRELACIÓN.

##### 4.1.2.2.1.- AUTOCORRELACIÓN GENERADA POR UN PROCESO AR(1) EN UN MODELO DE EFECTOS ALEATORIOS.

Dado el modelo  $y_{it} = \alpha + x'_{it}\beta + u_{it}$   $u_{it} = \mu_i + v_{it}$   $t = 1, 2, \dots, T$   $i = 1, 2, \dots, N$  donde

$$v_{it} = \rho v_{it-1} + \varepsilon_{it} \quad |\rho| < 1 \quad \varepsilon_{it} \sim IID(0, \sigma_\varepsilon^2)$$

Baltagi y Li (1991a) derivaron la transformación de Fuller y Battese (1974) correspondiente a este modelo.

Primeramente se aplica la transformación de Prais-Winsten:

$$C = \begin{bmatrix} (1-\rho^2)^{1/2} & 0 & 0 & \dots & 0 & 0 & 0 \\ -\rho & 1 & 0 & \dots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \dots & -\rho & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & -\rho & 1 \end{bmatrix}$$

Que nos sirve para transformar la perturbación aleatoria que sigue un proceso AR(1) en una perturbación clásica incorrelacionada.

Para datos de panel esta transformación tiene que aplicarse a cada uno de los N individuos del panel.

Vamos a estudiar cómo afecta esta transformación a la perturbación aleatoria. Sabemos que  $u_{it} = \mu_i + v_{it}$

$$u^* = (I_N \otimes C)u = (I_N \otimes C i_T)\mu + (I_N \otimes C)v$$

Ejemplo  $i=3; t=3$

$$u^* = (I_N \otimes C)u$$

$$\text{dimensión } I_N:3 \times 3 \quad C:3 \times 3 \quad (I_N \otimes C):9 \times 9 \quad u:9 \times 1 \quad (I_N \otimes C)u:9 \times 1$$

$$\begin{bmatrix} (1-\rho^2)^{1/2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\rho & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\rho & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (1-\rho^2)^{1/2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\rho & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\rho & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & (1-\rho^2)^{1/2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\rho & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\rho & 1 \end{bmatrix} \cdot \begin{bmatrix} u_{11} \\ u_{12} \\ u_{13} \\ u_{21} \\ u_{22} \\ u_{23} \\ u_{31} \\ u_{32} \\ u_{33} \end{bmatrix} = \begin{bmatrix} (1-\rho^2)^{1/2} u_{11} \\ -\rho u_{11} + u_{12} \\ -\rho u_{12} + u_{13} \\ (1-\rho^2)^{1/2} u_{21} \\ -\rho u_{21} + u_{22} \\ -\rho u_{22} + u_{23} \\ (1-\rho^2)^{1/2} u_{31} \\ -\rho u_{31} + u_{32} \\ -\rho u_{32} + u_{33} \end{bmatrix}$$

$(I_N \otimes C i_T) \mu$  donde  $C i_T = (1-\rho) i_T^\alpha$   $i_T^{\alpha'} = (\alpha, i_{T-1})$   $\alpha = \sqrt{\frac{1+\rho}{1-\rho}}$   $i_{T-1}$  = vector de unos de dimensión  $T-1$ .

Continuando con nuestro ejemplo

$$\left( \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \otimes (1-\rho) \begin{bmatrix} \sqrt{\frac{1+\rho}{1-\rho}} \\ 1 \\ 1 \end{bmatrix} \right) \cdot \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \left( \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} (1-\rho^2)^{1/2} \\ (1-\rho) \\ (1-\rho) \end{bmatrix} \right) \cdot \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} =$$

$$\begin{bmatrix} (1-\rho^2)^{1/2} & 0 & 0 \\ (1-\rho) & 0 & 0 \\ (1-\rho) & 0 & 0 \\ 0 & (1-\rho^2)^{1/2} & 0 \\ 0 & (1-\rho) & 0 \\ 0 & (1-\rho) & 0 \\ 0 & 0 & (1-\rho^2)^{1/2} \\ 0 & 0 & (1-\rho) \\ 0 & 0 & (1-\rho) \end{bmatrix} \cdot \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} (1-\rho^2)^{1/2} \mu_1 \\ (1-\rho) \mu_1 \\ (1-\rho) \mu_1 \\ (1-\rho^2)^{1/2} \mu_2 \\ (1-\rho) \mu_2 \\ (1-\rho) \mu_2 \\ (1-\rho^2)^{1/2} \mu_3 \\ (1-\rho) \mu_3 \\ (1-\rho) \mu_3 \end{bmatrix}$$

$$(I_N \otimes C)v$$

$$\left( \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} (1-\rho^2)^{1/2} & 0 & 0 \\ -\rho & 1 & 0 \\ 0 & -\rho & 1 \end{bmatrix} \right) \bullet \begin{bmatrix} v_{11} \\ v_{12} \\ v_{13} \\ v_{21} \\ v_{22} \\ v_{23} \\ v_{31} \\ v_{32} \\ v_{33} \end{bmatrix} =$$

$$= \begin{bmatrix} (1-\rho^2)^{1/2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\rho & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\rho & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (1-\rho^2)^{1/2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\rho & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\rho & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & (1-\rho^2)^{1/2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\rho & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\rho & 1 \end{bmatrix} \bullet \begin{bmatrix} v_{11} \\ v_{12} \\ v_{13} \\ v_{21} \\ v_{22} \\ v_{23} \\ v_{31} \\ v_{32} \\ v_{33} \end{bmatrix} = \begin{bmatrix} (1-\rho^2)^{1/2} v_{11} \\ -\rho v_{11} + v_{12} \\ -\rho v_{12} + v_{13} \\ (1-\rho^2)^{1/2} v_{21} \\ -\rho v_{21} + v_{22} \\ -\rho v_{22} + v_{23} \\ (1-\rho^2)^{1/2} v_{31} \\ -\rho v_{31} + v_{32} \\ -\rho v_{32} + v_{33} \end{bmatrix}$$

Por tanto la perturbación aleatoria transformada queda:

$$u^* = \begin{bmatrix} (1-\rho^2)^{1/2} \mu_1 \\ (1-\rho) \mu_1 \\ (1-\rho) \mu_1 \\ (1-\rho^2)^{1/2} \mu_2 \\ (1-\rho) \mu_2 \\ (1-\rho) \mu_2 \\ (1-\rho^2)^{1/2} \mu_3 \\ (1-\rho) \mu_3 \\ (1-\rho) \mu_3 \end{bmatrix} + \begin{bmatrix} (1-\rho^2)^{1/2} v_{11} \\ -\rho v_{11} + v_{12} \\ -\rho v_{12} + v_{13} \\ (1-\rho^2)^{1/2} v_{21} \\ -\rho v_{21} + v_{22} \\ -\rho v_{22} + v_{23} \\ (1-\rho^2)^{1/2} v_{31} \\ -\rho v_{31} + v_{32} \\ -\rho v_{32} + v_{33} \end{bmatrix}$$

Como  $Ci_T = (1-\rho)i_T^\alpha$   $i_T^{\alpha'} = (\alpha, i_{T-1})$   $\alpha = \sqrt{\frac{1+\rho}{1-\rho}}$  podemos reescribir la expresión

anterior de la siguiente forma:

$$u^* = (1-\rho)(I_N \otimes i_T^\alpha)\mu + (I_N \otimes C)v$$

Ahora calculamos la varianza de la perturbación aleatoria transformada:

$$\Omega^* = E[u^* u^{*'}] = \sigma_\mu^2 (1-\rho)^2 (I_N \otimes i_T^\alpha i_T^{\alpha'})\mu + (I_N \otimes C)E[vv'](I_N \otimes C')$$

Donde  $\sigma_\varepsilon^2 (I_N \otimes I_T) = (I_N \otimes C)E[vv'](I_N \otimes C')$  que sustituyéndolo en la expresión anterior nos queda

$$\Omega^* = \sigma_\mu^2 (1-\rho)^2 (I_N \otimes i_T^\alpha i_T^{\alpha'})\mu + \sigma_\varepsilon^2 (I_N \otimes I_T)$$

Utilizando el hecho de que  $d^2 = i_T^{\alpha'} i_T^\alpha = \alpha^2 + (T-1)$  esta expresión puede describirse como

$$\Omega^* = d^2 \sigma_\mu^2 (1-\rho)^2 \left( I_N \otimes \frac{i_T^\alpha i_T^{\alpha'}}{d^2} \right) \mu + \sigma_\varepsilon^2 (I_N \otimes I_T)$$

En el ejemplo que venimos utilizando  $d^2 = i_T^{\alpha'} i_T^\alpha = [\alpha \quad 1 \quad 1] \begin{bmatrix} \alpha \\ 1 \\ 1 \end{bmatrix} = \alpha^2 + (3-1)$

Llamamos  $J_T^\alpha = i_T^\alpha i_T^{\alpha'}$ , que en nuestro ejemplo es

$$J_T^\alpha = i_T^\alpha i_T^{\alpha'} = \begin{bmatrix} \alpha \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} \alpha & 1 & 1 \end{bmatrix} = \begin{bmatrix} \alpha^2 & \alpha & \alpha \\ \alpha & 1 & 1 \\ \alpha & 1 & 1 \end{bmatrix} \quad \text{como } d^2 \bar{J}_T^\alpha \text{ es idempotente donde}$$

$$\bar{J}_T^\alpha = \frac{i_T^\alpha i_T^{\alpha'}}{d^2} = \frac{J_T^\alpha}{d^2} \quad \text{que en nuestro ejemplo es } \bar{J}_T^\alpha = \begin{bmatrix} \alpha^2/d^2 & \alpha/d^2 & \alpha/d^2 \\ \alpha/d^2 & 1/d^2 & 1/d^2 \\ \alpha/d^2 & 1/d^2 & 1/d^2 \end{bmatrix}.$$

Siguiendo a Wansbeeck y Kapteyn (1982b), sustituimos  $I_T$  por  $E_T^\alpha + \bar{J}_T^\alpha$ , donde  $E_T^\alpha = I_T - \bar{J}_T^\alpha$ . Agrupando términos obtenemos la descomposición de la varianza.

$$\begin{aligned} \Omega^* &= d^2 \sigma_\mu^2 (1-\rho)^2 (I_N \otimes \bar{J}_T^\alpha) \mu + \sigma_\varepsilon^2 (I_N \otimes I_T) \\ \Omega^* &= d^2 \sigma_\mu^2 (1-\rho)^2 (I_N \otimes \bar{J}_T^\alpha) \mu + \sigma_\varepsilon^2 (I_N \otimes (E_T^\alpha + \bar{J}_T^\alpha)) \\ \Omega^* &= d^2 \sigma_\mu^2 (1-\rho)^2 (I_N \otimes \bar{J}_T^\alpha) \mu + \sigma_\varepsilon^2 (I_N \otimes \bar{J}_T^\alpha) + \sigma_\varepsilon^2 (I_N \otimes E_T^\alpha) \\ \Omega^* &= \left[ d^2 \sigma_\mu^2 (1-\rho)^2 \mu + \sigma_\varepsilon^2 \right] (I_N \otimes \bar{J}_T^\alpha) + \sigma_\varepsilon^2 (I_N \otimes E_T^\alpha) \end{aligned}$$

Llamando  $\sigma_\alpha^2 = \left[ d^2 \sigma_\mu^2 (1-\rho)^2 \mu + \sigma_\varepsilon^2 \right]$  y sustituyéndolo en la última expresión obtenemos la descomposición de la varianza.

$$\Omega^* = \sigma_\alpha^2 (I_N \otimes \bar{J}_T^\alpha) + \sigma_\varepsilon^2 (I_N \otimes E_T^\alpha)$$

Por otra parte

$$\sigma_\varepsilon \Omega^{*-1/2} = \frac{\sigma_\varepsilon}{\sigma_\alpha} (I_N \otimes \bar{J}_T^\alpha) + \frac{\sigma_\varepsilon}{\sigma_\varepsilon} (I_N \otimes E_T^\alpha) \quad \text{como } E_T^\alpha = I_T - \bar{J}_T^\alpha$$

$$\sigma_{\varepsilon} \Omega^{*-1/2} = \frac{\sigma_{\varepsilon}}{\sigma_{\alpha}} (I_N \otimes \bar{J}_T^{\alpha}) + \frac{\sigma_{\varepsilon}}{\sigma_{\varepsilon}} (I_N \otimes (I_T - \bar{J}_T^{\alpha}))$$

$$\sigma_{\varepsilon} \Omega^{*-1/2} = \frac{\sigma_{\varepsilon}}{\sigma_{\alpha}} (I_N \otimes \bar{J}_T^{\alpha}) + I_N \otimes I_T - I_N \otimes \bar{J}_T^{\alpha}$$

$$\sigma_{\varepsilon} \Omega^{*-1/2} = I_N \otimes I_T - \left[ 1 - \frac{\sigma_{\varepsilon}}{\sigma_{\alpha}} \right] (I_N \otimes \bar{J}_T^{\alpha})$$

Llamando  $\theta_{\alpha} = \left[ 1 - \frac{\sigma_{\varepsilon}}{\sigma_{\alpha}} \right]$  y sustituyendo en la expresión anterior nos queda

$$\sigma_{\varepsilon} \Omega^{*-1/2} = I_N \otimes I_T - \theta_{\alpha} (I_N \otimes \bar{J}_T^{\alpha})$$

A continuación se pre multiplican las observaciones transformadas de Prais-Winsten  $y^* = (I_N \otimes C)y$  por  $\sigma_{\varepsilon} \Omega^{*-1/2}$  obteniendo en ese caso  $y^{**} = \sigma_{\varepsilon} \Omega^{*-1/2} y^*$ , el típico elemento de  $y^{**}$  viene dado por:

$$y_{i1}^* - \theta_{\alpha} \alpha b_i, \quad y_{i2}^* - \theta_{\alpha} b_i, \quad \dots \quad y_{iT}^* - \theta_{\alpha} b_i$$

$$b_i = \alpha y_{i1}^* + \frac{\sum_{t=2}^T y_{it}^*}{d^2} \quad i = 1, 2, \dots, N$$

$$\text{Donde } \alpha = \sqrt{\frac{1+\rho}{1-\rho}} \quad d^2 = \alpha^2 + (T-1) \quad \sigma_{\alpha}^2 = d^2 \sigma_{\mu}^2 (1-\rho)^2 + \sigma_{\varepsilon}^2$$

$$\theta_{\alpha} = 1 - \frac{\sigma_{\varepsilon}}{\sigma_{\alpha}}$$

La primera observación recibe una atención especial en el modelo AR(1). La transformación de Prais-Winsten da un peso especial  $\sqrt{1-\rho^2}$  en  $y^*$ . Además la transformación de Fuller y Battese proporciona nuevamente un peso especial a

$\alpha = \sqrt{\frac{1+\rho}{1-\rho}}$  en el cálculo del promedio  $b_i$  y en la ecuación de pseudos-diferencias

$$y_{i1}^* - \theta_{\alpha} \alpha b_i, \quad y_{i2}^* - \theta_{\alpha} b_i, \quad \dots \quad y_{iT}^* - \theta_{\alpha} b_i.$$

Es preciso destacar que si  $\rho=0$ ,  $\alpha=1$ ,  $d^2=T$ ,  $\sigma_\alpha^2=\sigma_1^2$  y  $\theta_\alpha=\theta$ . El elemento de  $y_{it}^{**}$  se convierte en la familiar transformación  $(y_{it}-\theta\bar{y}_{i\cdot})$  del modelo sin autocorrelación.

Ahora bien ¿la transformación  $\theta$  propuesta por Baltagi coincide con la transformación  $\lambda$  propuesta por Wooldridge (2002; 287)?

Baltagi realiza la transformación del modelo del siguiente modo:

$$y_{it}-\theta\bar{y}_{i\cdot}, \quad \theta=1-\frac{\sigma_v}{\sigma_1}, \quad u_{it}=\mu+v_{it}, \quad V[v_{it}]=\sigma_v^2, \quad V[u_{it}]=\sigma_1^2=T\sigma_\mu^2+\sigma_v^2$$

Wooldridge propone la transformación siguiente:

$$y_{it}-\lambda\bar{y}_{i\cdot}, \quad \lambda=1-\left[\frac{1}{1+T\left(\frac{\sigma_c^2}{\sigma_u^2}\right)}\right]^{1/2}, \quad v_{it}=c+u_{it}, \quad \lambda=1-\left[\frac{1}{\frac{\sigma_u^2+T\sigma_c^2}{\sigma_u^2}}\right]^{1/2}$$

$\lambda=1-\left[\frac{\sigma_u^2}{\sigma_u^2+T\sigma_c^2}\right]^{1/2}$ , como  $\sigma_1^2=T\sigma_\mu^2+\sigma_v^2$  en la transformación de Baltagi y sustituyendo nos queda la transformación de Wooldridge del siguiente modo

$$\lambda=1-\left[\frac{\sigma_u^2}{\sigma_u^2+T\sigma_c^2}\right]^{1/2}, \quad \lambda=1-\left[\frac{\sigma_u^2}{\sigma_1^2}\right]^{1/2}, \quad \lambda=1-\frac{\sigma_u}{\sigma_1}, \quad \lambda=\theta$$

Los mejores estimadores cuadráticos insesgados (MECI) se derivan de la descomposición de  $\Omega^*$ . En efecto:

$$(I_N \otimes E_T^\alpha)u^* \square (0, \sigma_\varepsilon^2[I_N \otimes E_T^\alpha]) \quad y \\ (I_N \otimes J_T^\alpha)u^* \square (0, \sigma_\alpha^2[I_N \otimes \bar{J}_T^\alpha])$$



Por lo tanto

$$\hat{\sigma}_\varepsilon^2 = \frac{u^{*'} (I_N \otimes E_T^\alpha) u^*}{N(T-1)} \quad y \quad \hat{\sigma}_\alpha^2 = \frac{u^{*'} (I_N \otimes \bar{J}_T^\alpha) u^*}{N}$$

Para obtener los estimadores MECI, Baltagi y Li (1991a) sugieren estimar  $\rho$  a través

de los residuos WITHIN (la estimación de efectos fijos)  $\tilde{v}_{it}$  como  $\tilde{\rho} = \frac{\sum_{i=1}^N \sum_{t=2}^T \tilde{v}_{it} \tilde{v}_{it-1}}{\sum_{i=1}^N \sum_{t=2}^T \tilde{v}_{it-1}^2}$ . A

continuación se estiman  $\hat{\sigma}_\varepsilon^2$  y  $\hat{\sigma}_\alpha^2$  sustituyendo los residuos  $\hat{u}^*$  por los de la ecuación transformada de Prais-Winsten utilizando  $\tilde{\rho}$ .

Baltagi y Li (1997) encontraron que cuando  $T$  es pequeño (cosa que no sucede en nuestro caso) la estimación de  $\tilde{\rho}$  es bastante pobre y por esa razón recomendaron un estimador alternativo de  $\rho$  basado en la función de autocovarianzas  $Q_s = E[u_{it} u_{it-s}]$ .

Para el modelo AR(1),  $Q_s = \sigma_\mu^2 + \sigma_v^2 \rho^s$ . De  $Q_0, Q_1$  y  $Q_2$  es fácil demostrar que

$$\rho + 1 = \frac{Q_0 - Q_2}{Q_0 - Q_1}.$$

Un estimador consistente de  $\rho$ , cuando  $N$  es grande, viene dado por.

$$\hat{\rho} = \frac{\tilde{Q}_0 - \tilde{Q}_2}{\tilde{Q}_0 - \tilde{Q}_1} - 1 = \frac{\tilde{Q}_1 - \tilde{Q}_2}{\tilde{Q}_0 - \tilde{Q}_1} \quad \text{donde} \quad \tilde{Q}_s = \frac{\sum_{i=1}^N \sum_{t=s+1}^T \hat{u}_{it} \hat{u}_{it-s}}{N(T-S)}$$

Agrupados de la regresión  $y_{it} = \alpha + x'_{it} \beta + u_{it}$   $i = 1 \dots N$   $t = 1 \dots T$ .  $\hat{\sigma}_\varepsilon^2$  y  $\hat{\sigma}_\alpha^2$  se estiman sustituyendo los residuos MCO por los de la ecuación transformada de Prais Winsten usando  $\hat{\rho}$  en lugar de  $\tilde{\rho}$ .

Por tanto la estimación de un modelo con autocorrelación generada por un proceso AR(1) se realiza:

1. Aplicando la transformación de Prais Winsten en una primera etapa.
2. Restando un pseudo-promedio de los datos transformados:  
 $y_{i1}^* - \theta_\alpha \alpha b_i, \quad y_{i2}^* - \theta_\alpha b_i, \quad \dots \quad y_{iT}^* - \theta_\alpha b_i$  en la segunda etapa.

#### 4.1.2.2.2.- AUTOCORRELACIÓN GENERADA POR UN PROCESO AR(1) EN UN MODELO DE EFECTOS FIJOS.

Bhargava, Franzini y Narendranathan (1982) desarrollaron un procedimiento para estimar  $\beta$  en modelos de efectos fijos que presenten autocorrelación generada por un proceso AR(1).

Sea el modelo  $y_{it} = \alpha + x'_{it}\beta + u_{it}$   $u_{it} = \mu_i + v_{it}$   $t = 1, 2, \dots, T$   $i = 1, 2, \dots, N$  donde  $v_{it} = \rho v_{it-1} + \varepsilon_{it}$   $|\rho| < 1$   $\varepsilon_{it} \square IID(o, \sigma_\varepsilon^2)$ , el estimador de Mínimos Cuadrados Generalizados (MCG), en caso de autocorrelación, se define como:

$$b = \left( X^{*'} \Omega^{*-1} X^* \right)^{-1} X^{*'} \Omega^{*-1} y$$

Donde  $\Omega^* = (I_N \otimes \Omega)$  es una matriz de rango  $NT \times NT$  en la que  $\Omega$  es la matriz habitual de  $T \times T$  de varianzas y covarianzas de un proceso AR(1) estacionario, en el que los elementos de  $\Omega$  tienen la forma  $w_{tk} = \frac{\rho^{|t-k|}}{1-\rho^2}$

Desde el punto de vista de los cálculos, la expresión  $b = \left( X^{*'} \Omega^{*-1} X^* \right)^{-1} X^{*'} \Omega^{*-1} y$  resulta poco adecuada y las estimaciones MCG pueden obtenerse del siguiente modo.

En primer lugar consideramos el modelo siguiente:

$$y_{it} = \delta_i + \sum_{j=1}^h x_{ijt} \gamma_j + v_{it} \quad t = 1, \dots, T; i = 1, \dots, N$$

$$v_{it} = \rho v_{it-1} + \varepsilon_{it} \quad t = 1, \dots, T; i = 1, \dots, N$$

Posteriormente, lo transformamos siguiendo la transformación habitual en presencia de autocorrelación

$$\sqrt{1-\rho^2} y_{i1} = \sqrt{1-\rho^2} \delta_i + \sqrt{1-\rho^2} \sum_{j=1}^h x_{ijt} \gamma_j + \varepsilon_{it} \quad t=1, \dots, T; i=1, \dots, N$$

y

$$(y_{it} - \rho y_{it-1}) = (1-\rho) \delta_i + \sum_{j=1}^h (x_{ijt} - \rho x_{ijt-1}) \gamma_j + \varepsilon_{it} \quad t=2, \dots, T; i=1, \dots, N$$

A continuación eliminamos la variable ficticia  $\delta_i$  utilizando la matriz simétrica idempotente  $C^+ = (I_H \otimes C^*)$  donde  $C^* = I_T - w(w'w)^{-1}w'$  y  $w$  es un vector  $T \times 1$  dado por  $w' = (\sqrt{1-\rho^2}, 1-\rho, \dots, 1-\rho) = \left( \frac{\sqrt{1+\rho}}{\sqrt{1-\rho}}, 1, \dots, 1 \right) (1-\rho)$ .

Es fácil demostrar que esta transformación es equivalente a transformar la variable dependiente como

$$\sqrt{1-\rho^2} \left( \frac{(T-1)(1-\rho)}{T(1-\rho)+2\rho} \right) y_{i1} - \frac{\sqrt{1-\rho^2}}{T(1-\rho)+2\rho} \sum_{t=2}^T (y_{it} - \rho y_{it-1}) \quad i=1, \dots, N$$

Y

$$(y_{it} - \rho y_{it-1}) - \frac{(1-\rho^2) y_{i1}}{T(1-\rho)+2\rho} - \frac{(1-\rho)}{T(1-\rho)+2\rho} \sum_{t=2}^T (y_{it} - \rho y_{it-1})$$

$i=1, \dots, N \quad t=2, \dots, T$

Las X pueden transformarse de forma parecida y a continuación correr una regresión MCO sobre las Y y X transformadas que producirá estimaciones eficientes de los parámetros.

La discusión anterior asume que  $\rho$  es conocido. En la práctica, sin embargo,  $\rho$  será desconocido y, tal y como afirmó Kiefer (1980) “las estimaciones máximo verosímiles pueden no poseer las propiedades asintóticas habituales dada la incidencia de los parámetros”; por lo tanto no será posible estimar  $\rho$  de forma consistente por ningún tipo de procedimiento de máxima verosimilitud a no ser que el número de periodos

temporales  $T$  sea muy grande. La última condición es poco probable que se cumpla en la práctica, por lo que sería preciso desarrollar un procedimiento específico para AR(1).

Suponiendo que tenemos una muestra grande en  $N$  (pero no en  $T$ ) escribimos

$$E[d_p] = E \left[ \frac{\sum_{i=1}^N \sum_{t=2}^T (v_{it} - v_{it-1})^2}{\sum_{i=1}^N \sum_{t=1}^T (v_{it} - v_{i\Box})^2} \right] \text{ donde } v_{i\Box} = \frac{1}{T} \sum_{t=1}^T v_{it} \quad (i=1 \dots N) \text{ y las } v_{it} \text{ son los residuos}$$

MCO de la estimación  $y_{it} = \delta_i + \sum_{j=1}^h x_{ijt} \gamma_j + v_{it} \quad t=1, \dots, T; i=1, \dots, N$ . Ahora utilizando la aproximación de Nagar (1959) para  $E[d_p]$  y teniendo en cuenta que todos los términos de sección cruzada están divididos por  $N$  (y por tanto pueden ser ignorados) tenemos

$$E[d_p] = \frac{E \left[ \frac{1}{NT} \sum_{i=1}^N \sum_{t=2}^T (v_{it} - v_{it-1})^2 \right]}{E \left[ \frac{1}{NT} \sum_{i=1}^N \sum_{t=2}^T (v_{it} - v_{it-1})^2 \right]} \text{ Tomando esperanzas dentro de los paréntesis, al}$$

cuadrado y eliminando  $\sigma^2$  tenemos  $E[d_p] = \frac{2(1-\rho)(T-1)/T}{\left[ 1 - \left( \frac{1}{T^2} \right) \sum_{s=1}^T \sum_{t=1}^T \rho^{|s-t|} \right]}$ . Si definimos, de

forma parecida al análisis de series temporales, una estimación de  $\rho_d$  (que es consistente cuando  $T \rightarrow \infty$ ) por la expresión:  $\rho_d = 1 - \frac{d_p}{2}$

Entonces  $E[\rho_d] = 1 - \frac{(1-\rho)(T-1)}{\left[ T - \left( \frac{1}{T} \right) \sum_{s=1}^T \sum_{t=1}^T \rho^{|s-t|} \right]}$  o bien la expresión alternativa

$$E[\rho_d] = 1 - \frac{(1-\rho)(T-1)}{\left[ T - \frac{1+\rho}{1-\rho} + \frac{2\rho(1-\rho^T)}{T(1-\rho)^2} \right]}$$

Es preciso remarcar que como  $\rho_d$  será calculado siempre para testar si los errores presentan autocorrelación,  $\rho_D$  puede calcularse fácilmente de la expresión  $\rho_d = 1 - \frac{d_p}{2}$  y sustituir este valor en el lado izquierdo de las expresiones

$$E[\rho_D] = 1 - \frac{(1-\rho)(T-1)}{\left[ T - \left( \frac{1}{T} \right) \sum_{s=1}^T \sum_{t=1}^T \rho^{|s-t|} \right]} \quad \text{o} \quad E[\rho_D] = 1 - \frac{(1-\rho)(T-1)}{\left[ T - \frac{1+\rho}{1-\rho} + \frac{2\rho(1-\rho^T)}{T(1-\rho)^2} \right]}, \quad \text{una buena}$$

estimación de  $\rho$  (sea  $\rho_D^*$ ) puede obtenerse resolviendo la ecuación para  $\rho$  de forma iterativa.

## 4.2.- UN PLANTEAMIENTO ALTERNATIVO.

Cada vez resulta más frecuente, especialmente en los trabajos de carácter práctico, utilizar estimadores que sean robustos a, bien a heteroscedasticidad, bien a autocorrelación e incluso robustos a la presencia de ambos problemas.

La ventaja resulta evidente, no tenemos que preocuparnos de como corregir los problemas, y nos centramos exclusivamente en la significatividad de las variables introducidas en el modelo, aunque si bien es cierto, siempre a costa de perder precisión.

En datos de panel, como no podía ser de otra manera, también están disponibles estimadores de este tipo. Más concretamente, en este trabajo se han utilizado “cluster robust standard error” (errores robustos por cluster o conglomerado).

Carter Hill (2012:85) nos muestra, a través de un modelo de datos de panel muy sencillo, el funcionamiento de estos errores robustos por cluster.

Primeramente plantea el método en un modelo de corte transversal, para posteriormente aplicarlo a datos de panel.

Sea el modelo de corte transversal  $y_i = \beta_1 + \beta_2 x_i + e_i$ , en este modelo se demuestra

que  $\hat{\beta}_2 = \beta_2 + \sum w_i e_i$

$$\hat{\beta}_2 = \sum w_i y_i = \sum w_i (\beta_1 + \beta_2 x_i + e_i)$$

$$\hat{\beta}_2 = \sum w_i \beta_1 + \beta_2 \sum w_i x_i + \sum w_i e_i$$

$$\text{Como resulta que } \sum w_i = \sum \left[ \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \right] = \frac{\sum (x_i - \bar{x})}{\sum (x_i - \bar{x})^2} = \frac{0}{\sum (x_i - \bar{x})^2} = 0$$

Trabajamos a continuación con la expresión siguiente:

$$\sum w_i x_i = \sum \left[ \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \right] x_i = \frac{\sum (x_i - \bar{x}) x_i}{\sum (x_i - \bar{x})^2}$$

Desarrollando el numerador  $\sum (x_i - \bar{x}) x_i = \sum x_i^2 - \bar{x} \sum x_i = \sum x_i^2 - n\bar{x}$

Haciendo lo mismo con el denominador

$$\begin{aligned} \sum (x_i - \bar{x})^2 &= \sum x_i^2 - 2\bar{x} \sum x_i + n\bar{x}^2 \\ \sum (x_i - \bar{x})^2 &= \sum x_i^2 - 2n\bar{x}^2 + n\bar{x}^2 \\ \sum (x_i - \bar{x})^2 &= \sum x_i^2 - n\bar{x}^2 \end{aligned}$$

Como numerador y denominador son iguales se concluye que  $\sum w_i x_i$  es igual a 1.

Por lo tanto la expresión:  $\hat{\beta}_2 = \beta_1 \sum_0 w_i + \beta_2 \underbrace{\sum w_i x_i}_1 + \sum w_i e_i$

Se simplifica en  $\hat{\beta}_2 = \beta_2 + \sum w_i e_i$

El siguiente paso es calcular la varianza del estimador:  $Var(\hat{\beta}_2) = E[\hat{\beta}_2 - E(\hat{\beta}_2)]^2$

Como el estimador de Mínimos Cuadrados Ordinarios es un estimador insesgado, utilizando esta propiedad y sustituyendo el estimador por su expresión equivalente llegamos a:

$$\begin{aligned} Var(\hat{\beta}_2) &= E[\beta_2 + \sum w_i e_i - \beta_2]^2 = E[\sum w_i e_i]^2 = E\left[\sum w_i^2 e_i^2 + 2 \sum_{i \neq j} w_i w_j e_i e_j\right] \\ Var(\hat{\beta}_2) &= \sum w_i^2 \underbrace{E[e_i^2]}_{\sigma^2} + 2 \sum_{i \neq j} w_i w_j \underbrace{E[e_i e_j]}_0 \\ Var(\hat{\beta}_2) &= \sigma^2 \sum w_i^2 = \sigma^2 \sum \left[ \frac{(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \right]^2 = \sigma^2 \frac{\sum (x_i - \bar{x})^2}{\sum (x_i - \bar{x})^4} = \frac{\sigma^2}{\sum (x_i - \bar{x})^2} \end{aligned}$$

La varianza del estimador toma esta expresión *siempre* que se cumplan las hipótesis de homoscedasticidad y no autocorrelación. Ahora bien, ¿qué sucede si no se cumple la hipótesis de homoscedasticidad?



$$\begin{aligned}
Var(\hat{\beta}_2) &= Var\left(\sum w_i e_i\right) \\
Var(\hat{\beta}_2) &= \sum w_i^2 Var(e_i) + \sum_{i \neq j} \sum w_i w_j \underbrace{cov(e_i e_j)}_0 \\
Var(\hat{\beta}_2) &= \frac{\sum (x_i - \bar{x})^2}{\sum (x_i - \bar{x})^4} \sigma_i^2
\end{aligned}$$

La varianza es ahora, distinta para cada individuo.

Pasamos a continuación al modelo de datos de panel. El objetivo es aplicar estos conceptos a un modelo diferente.

$$y_{it} = \beta_1 + \beta_2 x_{it} + e_{it}$$

Con las hipótesis:

$cov(e_{it} e_{is}) = \varphi_{ts}$  Varianza constante en los individuos en distintos momentos del tiempo.

$cov(e_{it} e_{js}) = 0; i \neq j$  No existe relación entre los individuos en diferentes momentos del tiempo.

El estimador de MCO Agrupados del parámetro  $\beta_2$  viene dado por la expresión:

$$\hat{\beta}_2 = \beta_2 + \sum_{i=1}^N \sum_{t=1}^T w_{it} e_{it}$$

Donde  $w_{it} = \frac{x_{it} - \bar{\bar{x}}}{\sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{\bar{x}})^2}$  y  $\bar{\bar{x}} = \frac{\sum_{i=1}^N \sum_{t=1}^T x_{it}}{NT}$  por lo tanto.

$$Var(\hat{\beta}_2) = Var\left(\sum_{i=1}^N \sum_{t=1}^T w_{it} e_{it}\right) = Var\left(\sum_{i=1}^N g_i\right)$$

Como los errores de los diferentes individuos están incorrelacionados, esto significa que  $g_i$  está incorrelacionado con  $g_j$  para  $i \neq j$ . Por lo tanto la varianza del estimador queda como:

$$Var(\hat{\beta}_2) = Var\left(\sum_{i=1}^N g_i\right) = \sum_{i=1}^N Var(g_i) + \sum_{i=1}^N \sum_{j=i+1}^N 2\text{cov}(g_i g_j) = \sum_{i=1}^N Var(g_i)$$

#### 4.3.- TEST DE ESPECIFICACIÓN DE HAUSMAN.

Llegado a este punto, se plantea una cuestión primordial ¿con qué modelo debemos trabajar, con el de Efectos Fijos o con el de Efectos Aleatorios?

Responder a esta pregunta supone comprobar si se cumple una hipótesis básica en el modelo de componentes del error. Esto es, si se cumple que  $E[u_{it}/X_{it}] = 0$ . La importancia del cumplimiento de esta hipótesis viene dada porque la perturbación aleatoria contiene efectos invariantes individuales (los  $\mu_i$ ) que son efectos no observables y que pueden estar correlacionados con las  $X_{it}$ .

Si  $E[u_{it}/X_{it}] \neq 0$  el estimador  $\hat{\beta}_{MCG}$  que usamos en el modelo de Efectos Aleatorios se convierte en un estimador SESGADO e INCONSISTENTE de  $\beta$ . Por otra parte, con la transformación Within eliminamos los  $\mu_i$  y el estimador  $\hat{\beta}_{Within}$  es un estimador SESGADO y CONSISTENTE de  $\beta$ . Para resolver esta cuestión Hausman (1978) sugirió comparar ambos estimadores:  $\hat{\beta}_{MCG}$  y  $\hat{\beta}_{Within}$ .

Bajo la hipótesis nula  $H_0: E[u_{it}/X_{it}] = 0$  ambos estimadores son consistentes, en el caso de que  $H_0$  sea falsa tendrán límites en probabilidad diferentes. Esto es, el estimador  $\hat{\beta}_{Within}$  es consistente tanto si  $H_0$  es verdadera o falsa, mientras que el estimador  $\hat{\beta}_{MCG}$  es ELIO, consistente y asintóticamente eficiente bajo  $H_0$ , pero es INCONSISTENTE cuando  $H_0$  es falsa.

Un contraste estadístico natural debería basarse en comparar ambos estimadores:

$$\tilde{q}_1 = \hat{\beta}_{MCG} - \hat{\beta}_{Within} \quad \text{Bajo } H_0, \quad p \lim \hat{q}_1 = 0 \text{ y } \text{cov}(\hat{q}_1, \hat{\beta}_{MCG}) = 0.$$

Para demostrarlo, vamos a usar el hecho de que  $\hat{\beta}_{MCG} - \beta = (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} u$

$$\begin{aligned}
\hat{\beta}_{MCG} &= (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} y \\
\hat{\beta}_{MCG} &= (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} (X \beta + u) \\
\hat{\beta}_{MCG} &= (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} X \beta + (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} u \\
\hat{\beta}_{MCG} &= \beta + (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} u \\
\hat{\beta}_{MCG} - \beta &= (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} u
\end{aligned}$$

y que  $\hat{\beta}_{Within} - \beta = (X' Q^{-1} X)^{-1} X' Q^{-1} u$

$$\begin{aligned}
\hat{\beta}_{Within} &= (X' Q^{-1} X)^{-1} X' Q^{-1} y \\
\hat{\beta}_{Within} &= (X' Q^{-1} X)^{-1} X' Q^{-1} (X \beta + u) \\
\hat{\beta}_{Within} &= (X' Q^{-1} X)^{-1} X' Q^{-1} X \beta + (X' Q^{-1} X)^{-1} X' Q^{-1} u \\
\hat{\beta}_{Within} &= \beta + (X' Q^{-1} X)^{-1} X' Q^{-1} u \\
\hat{\beta}_{Within} - \beta &= (X' Q^{-1} X)^{-1} X' Q^{-1} u
\end{aligned}$$

De aquí llegamos a que  $E(\hat{q}_1) = 0$

$$\begin{aligned}
\hat{q}_1 &= \hat{\beta}_{MCG} - \hat{\beta}_{Within} \\
\hat{q}_1 &= \beta + (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} u - \left( \beta + (X' Q^{-1} X)^{-1} X' Q^{-1} u \right) \\
\hat{q}_1 &= (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} u - (X' Q^{-1} X)^{-1} X' Q^{-1} u
\end{aligned}$$

Tomando esperanzas nos queda

$$\begin{aligned}
E(\hat{q}_1) &= (X' \Omega^{-1} X)^{-1} X' \Omega^{-1} E(u) - (X' Q^{-1} X)^{-1} X' Q^{-1} E(u) \text{ como } E(u) = 0 \text{ sustituyendo} \\
E(\hat{q}_1) &= 0
\end{aligned}$$

Ahora demostraremos que  $\text{cov}(\hat{q}_1, \hat{\beta}_{MCG}) = 0$

$$\begin{aligned}
\text{cov}(\hat{\beta}_{MCG}, \hat{q}_1) &= \text{Var}[\hat{\beta}_{MCG}] - \text{cov}[\hat{\beta}_{MCG}, \hat{\beta}_{Within}] = 0 \\
\text{cov}(\hat{\beta}_{MCG}, \hat{q}_1) &= E\left[\left(\hat{\beta}_{MCG} - \underbrace{E(\hat{\beta}_{MCG})}_{\beta}\right)\left(\hat{q}_1 - \underbrace{E(\hat{q}_1)}_0\right)\right] = E\left[\left(\hat{\beta}_{MCG} - \beta\right)\hat{q}_1\right] \\
\text{cov}(\hat{\beta}_{MCG}, \hat{q}_1) &= E\left[\left(\hat{\beta}_{MCG} - \beta\right)\left(\hat{\beta}_{MCG} - \hat{\beta}_{Within}\right)\right]; \\
\text{cov}(\hat{\beta}_{MCG}, \hat{q}_1) &= E[\hat{\beta}_{MCG}]^2 - E[\hat{\beta}_{MCG}, \hat{\beta}_{Within}] - \underbrace{\beta E[\hat{\beta}_{MCG}]}_{\beta} + \underbrace{\beta E[\hat{\beta}_{Within}]}_{\beta} \\
\text{cov}(\hat{\beta}_{MCG}, \hat{q}_1) &= E[\hat{\beta}_{MCG}]^2 - E[\hat{\beta}_{MCG}, \hat{\beta}_{Within}] - \beta^2 + \beta^2
\end{aligned}$$

Reordenando tenemos.

$$\text{cov}(\hat{\beta}_{MCG}, \hat{q}_1) = \left(E[\hat{\beta}_{MCG}]^2 - \beta^2\right) - \left(E[\hat{\beta}_{MCG}, \hat{\beta}_{Within}] - \beta^2\right)$$

Recordemos que  $\text{Var}[\hat{\beta}_{MCG}] = E[\hat{\beta}_{MCG} - E[\hat{\beta}_{MCG}]]^2 = E[\hat{\beta}_{MCG} - \beta]^2$  deshaciendo el cuadrado llegamos a

$$\text{Var}[\hat{\beta}_{MCG}] = E[\hat{\beta}_{MCG}]^2 - 2\beta E[\hat{\beta}_{MCG}] + \beta^2 = E[\hat{\beta}_{MCG}]^2 - 2\beta^2 + \beta^2 = E[\hat{\beta}_{MCG}]^2 - \beta^2$$

Por lo tanto en la expresión de la covarianza podemos sustituir el valor de  $\left(E[\hat{\beta}_{MCG}]^2 - \beta^2\right)$  por su expresión equivalente:  $\text{Var}[\hat{\beta}_{MCG}]$

Repetimos el mismo camino para  $\text{cov}[\hat{\beta}_{MCG}, \hat{\beta}_{Within}]$  y obtenemos una expresión

equivalente:

$$\begin{aligned}
\text{cov}[\hat{\beta}_{MCG}, \hat{\beta}_{Within}] &= E\left[\left(\hat{\beta}_{MCG} - \underbrace{E[\hat{\beta}_{MCG}]}_{\beta}\right)\left(\hat{\beta}_{Within} - \underbrace{E[\hat{\beta}_{Within}]}_{\beta}\right)\right] \\
\text{cov}[\hat{\beta}_{MCG}, \hat{\beta}_{Within}] &= E\left[\left(\hat{\beta}_{MCG} - \beta\right)\left(\hat{\beta}_{Within} - \beta\right)\right]
\end{aligned}$$

Calculando las esperanzas de los productos cruzados obtenemos:

$$\begin{aligned}
\text{cov}[\hat{\beta}_{MCG}, \hat{\beta}_{Within}] &= E[\hat{\beta}_{MCG}, \hat{\beta}_{Within}] - \beta E[\hat{\beta}_{Within}] - \beta E[\hat{\beta}_{MCG}] + \beta^2 \\
\text{cov}[\hat{\beta}_{MCG}, \hat{\beta}_{Within}] &= E[\hat{\beta}_{MCG}, \hat{\beta}_{Within}] - \beta^2 - \beta^2 + \beta^2
\end{aligned}$$

Y llegamos a afirmar que  $\text{cov}[\hat{\beta}_{MCG}, \hat{\beta}_{Within}] = E[\hat{\beta}_{MCG}, \hat{\beta}_{Within}] - \beta^2$ , por tanto queda demostrado que  $\text{cov}(\hat{\beta}_{MCG}, \hat{q}_1) = \text{Var}[\hat{\beta}_{MCG}] - \text{cov}[\hat{\beta}_{MCG}, \tilde{\beta}_{WITHIN}]$

Llegados a este punto lo único que nos queda es sustituir  $\text{Var}[\hat{\beta}_{MCG}]$  y  $\text{cov}[\hat{\beta}_{MCG}, \hat{\beta}_{Within}]$  por su valor. Siguiendo a Baltagi (2001:65)

$$\text{cov}(\hat{\beta}_{MCG}, \hat{q}_1) = (X' \Omega^{-1} X)^{-1} - (X' \Omega^{-1} X)^{-1} X \Omega^{-1} E(uu') QX (X' QX)^{-1}$$

$$\text{cov}(\hat{\beta}_{MCG}, \hat{q}_1) = (X' \Omega^{-1} X)^{-1} - (X' \Omega^{-1} X)^{-1} = 0$$

Utilizando el hecho de que  $\hat{q}_1 = \hat{\beta}_{MCG} - \hat{\beta}_{Within}$  y por tanto  $\hat{\beta}_{Within} = \hat{\beta}_{MCG} - \hat{q}_1$  tomando varianzas tendremos  $\text{Var}(\hat{\beta}_{Within}) = \text{Var}(\hat{\beta}_{MCG}) + \text{Var}(\hat{q}_1) - \underbrace{2\text{cov}(\hat{\beta}_{MCG}, \hat{q}_1)}_0$ . Por lo tanto

$$\text{Var}(\hat{q}_1) = \text{Var}(\hat{\beta}_{Within}) - \text{Var}(\hat{\beta}_{MCG}) = \sigma_v^2 (X' QX)^{-1} - (X' \Omega^{-1} X)^{-1}.$$

En consecuencia el estadístico de Hausman viene determinado por:  $m_1 = \hat{q}_1' [\text{Var}(\hat{q}_1)]^{-1} \hat{q}_1$  que bajo  $H_0$  se distribuye asintóticamente como una  $\chi_K^2$  donde  $K$  denota la dimensión del vector  $\beta$ .

Para hacer este contraste operacional, es preciso reemplazar  $\Omega$  por un estimador consistente  $\hat{\Omega}$ , y la estimación MCG por su correspondiente estimación MCG FACTIBLE.

#### 4.4.- TEST DE AGRUPABILIDAD.

Así como el test de Hausman tiene sentido para elegir entre la estimación de efectos fijos frente a la de efectos aleatorios, los contrastes que planteamos a continuación tienen un campo de aplicación más general. Este campo de aplicación es conocer si la modelización mediante de datos de panel tiene sentido.

Si seguimos a Hsiao (2002: 13) estos contrastes plantean tres modelos distintos a aplicar con la intención de elegir el modelo más adecuado a nuestros datos.

$$y_{it} = \alpha_i + \beta_i' x_{it} + u_{it}$$

Este modelo supone que cada individuo tiene  $\alpha$  y  $\beta$  distinta. Estamos en el caso de un modelo para cada individuo que es el caso menos restrictivo posible.

Llamamos  $S_1$  a la suma de los residuos al cuadrado. Para calcularlo hacemos una regresión para cada individuo y sumamos los residuos al cuadrado de todas y cada una de las regresiones realizadas (una por individuo)

$$y_{it} = \alpha_i + \beta' x_{it} + u_{it}$$

Este modelo contempla un  $\alpha$  distinto para cada individuo. Se trata del modelo de Efectos Individuales Específicos que puede ser estimado por los métodos de Efectos Fijos, Efectos Aleatorios, Primera Diferencia, etc.

Generalmente se suele estimar por Efectos Fijos (para elegir entre Efectos Fijos o Aleatorios ya tenemos el test de Hausman). Una vez estimado calculamos la suma de los residuos al cuadrado. A esta suma la vamos a llamar  $S_2$

$$y_{it} = \alpha + \beta' x_{it} + u_{it}$$

Este modelo supone establecer la restricción de que todos los parámetros son iguales para cada individuo. Es el modelo MCO AGRUPADO que es el modelo más restrictivo posible.

Llamamos  $S_3$  a la suma de los residuos al cuadrado de la regresión MCO agrupado.

Como puede observarse, los contrastes planteados son contraste de comparación de suma de residuos al cuadrado, propuestos por Chow (1960) y utilizados muy habitualmente en otras áreas de la econometría. Los estadísticos de contraste resultantes son estadístico F ya que los contrastes plantean un cociente de sumas de cuadrados y cada una de ellas sigue una chi cuadrado.



## Hipótesis a plantear

## Estadístico de Contraste

$$H_3 : \alpha_1 = \alpha_2 = \dots \alpha_N$$
$$\beta_1 = \beta_2 = \dots \beta_N$$

$$F_3 = \frac{\frac{S_3 - S_1}{(N-1)(K+1)}}{\frac{S_1}{NT - N(K+1)}}$$

Regla de contrastación: Rechazar  $H_3$  si

$$F_3 \geq F_{(N-1)(K+1), NT-N(K+1)}^{(\alpha)}$$

Si aceptamos  $H_3$  significa que el modelo

adecuado es  $y_{it} = \alpha + \beta' x_{it} + u_{it}$

Rechazar  $H_3$  supone admitir que este modelo no es el adecuado y por tanto, la heterogeneidad estará en el intercepto o en la pendiente.

$$H_1 : \beta_1 = \beta_2 = \dots \beta_N$$

$$F_1 = \frac{\frac{S_2 - S_1}{(N-1)K}}{\frac{S_1}{NT - N(K+1)}}$$

Regla de contrastación: Rechazar  $H_1$  si

$$F_1 \geq F_{(N-1)K, NT-N(K+1)}^{(\alpha)}$$

Si aceptamos  $H_1$  significa que el modelo

adecuado es  $y_{it} = \alpha_i + \beta' x_{it} + u_{it}$

Si rechazamos  $H_1$  tenemos un modelo para cada

individuo.

En el caso de aceptar  $H_1$  podemos aplicar un contraste condicionado para interceptos homogéneos

$$H_4 : \alpha_1 = \alpha_2 = \dots \alpha_N$$

$$\text{dado } \beta_1 = \beta_2 = \dots \beta_N$$

$$F_4 = \frac{\frac{S_3 - S_2}{(N-1)}}{\frac{S_2}{N(T-1) - K}}$$

Regla de contrastación: Rechazar  $H_4$  si

$$F_4 \geq F_{(N-1), N(T-1)-K}^{(\alpha)}$$

Este contraste tiene sentido aplicarlo para elegir entre el modelo  $y_{it} = \alpha_i + \beta' x_{it} + u_{it}$  de efectos individuales específicos y el modelo agrupado  $y_{it} = \alpha + \beta' x_{it} + u_{it}$ .

Rechazar la hipótesis nula significa que el modelo adecuado es el de efectos individuales específicos.

Con estos contrastes se pretende determinar qué modelo de datos de panel utilizar, así, si concluimos que los parámetros son iguales para todos los individuos, el modelo idóneo es el Datos de Panel Agrupados. Ahora bien, si esta hipótesis resulta rechaza, se deberá decir si el modelo idóneo es el que tiene un intercepto común y pendiente distinta para los distintos individuos o el modelo más general que presenta intercepto y pendiente distinta para cada uno de los individuos.

#### 4.5.- CONTRASTE RESET (Regression Specification Error Test) O DE ERROR DE ESPECIFICACIÓN.

Carter Hill, Griffiths y Lim (2012:151) hacen referencia en su obra, al contraste RESET de Ramsey (1969). Este contraste fue propuesto por Ramsey para detectar variables omitidas y formas funcionales incorrectas. El autor, en el artículo original lo compara además con otros contrastes alternativos.

Con datos transversales su funcionamiento es el siguiente.

En un primer paso se estima el modelo y se obtienen los valores ajustados del modelo.

$$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + u_i$$
$$\hat{y}_i = b_1 + b_2 x_{i2} + b_3 x_{i3}$$

En un segundo momento, se considera los dos modelos artificiales siguientes:

$$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \gamma_1 \hat{y}_i^2 + u_i$$
$$y_i = \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \gamma_1 \hat{y}_i^2 + \gamma_2 \hat{y}_i^3 + u_i$$

Un contraste de error de especificación es testar si  $H_0 : \gamma_1 = 0$  frente a la hipótesis alternativa  $H_1 : \gamma_1 \neq 0$  en el primer modelo. En el segundo modelo se contrasta la hipótesis:  $H_1 : \gamma_1 = \gamma_2 = 0$  frente a  $H_1 : \gamma_1 \neq 0$  y/o  $\gamma_2 \neq 0$ . En el primer caso podemos utilizar un contraste  $t$  o  $F$ . En el segundo caso utilizaremos un test  $F$ .

Rechazar la hipótesis nula implica que el modelo original no es adecuado y puede ser mejorado. Recordemos que  $\hat{y}_i^2$  y  $\hat{y}_i^3$  son funciones polinomiales de  $x_{i2}$  y  $x_{i3}$  (como por ejemplo  $x_{i2}^2, x_{i3}^3, x_{i2}x_{i3}, x_{i2}x_{i3}^2$ ) Estos polinomios pueden aproximar muchos tipos de formas funcionales, si la forma funcional original no es correcta la aproximación polinomial que incluya  $\hat{y}_i^2$  y  $\hat{y}_i^3$  puede mejorar significativamente el ajuste del modelo. No sólo eso, si hemos omitido variables, y esas variables están

correlacionadas con  $x_{i2}$  y  $x_{i3}$ , entonces también lo estarán con sus cuadrados y alguno de sus efectos pueden ser recogidos al incluir los términos  $\hat{y}_i^2$  y  $\hat{y}_i^3$ .

La filosofía general del test es: si podemos mejorar significativamente el modelo aumentando artificialmente su poder de predicción, entonces el modelo original no es adecuado.

Este contraste, utilizado en el contexto de datos de panel, tiene el funcionamiento siguiente:

1. Estimamos el modelo con el estimador de datos de panel oportuno; puede ser Efectos Fijos, Efectos Aleatorios, etc. En este trabajo el contraste se ha aplicado estimando el modelo por Efectos Fijos.
2. Una vez estimado el modelo, guardamos la variable estimada al cuadrado y al cubo.
3. Volvemos a estimar el modelo por Efectos Fijos, pero ahora se incluye como variable explicativa las anteriores más la variable estimada al cuadrado.
4. Sobre el nuevo modelo, realizamos un contraste individual sobre la nueva variable introducida. La nueva variable puede ser significativa (significa que alguna combinación de dimensión 2 de las variables originales hemos omitido y por tanto hay que buscarla) o no significativa. Si este es el caso, pasamos a la fase siguiente.
5. Estimamos el modelo original añadiéndoles como variables explicativas la variable estimada al cuadrado y al cubo.
6. Sobre este nuevo modelo, volvemos a realizar un contraste conjunto de significación de las dos nuevas variables introducidas. Si las variables no son significativas, la forma funcional original es correcta. Si las variables son

significativas nos proporciona información sobre las variables que hemos omitido y que debemos introducir en el modelo.

7. El proceso termina cuando sobre el modelo reformulado, las variables estimadas al cuadrado y al cubo resultan ser no significativas.

## **5.- ESPECIFICACIÓN DEL MODELO.**

En este capítulo se trata de abordar las dos primeras hipótesis que nos proponíamos contestar en esta tesis.

La primera ¿Es posible construir un modelo que explique el comportamiento comercial de las marcas en el punto de venta? Esta pregunta será contestada en los apartados siguientes, donde se especificarán diferentes modelos hasta por fin quedarnos con un único modelo o más bien, un método de estimación del modelo.

La segunda hipótesis, estaba relacionada con la primera; suponiendo que podamos construir el modelo, ¿este modelo es único? o por el contrario será necesario construir tantos modelos como formatos y establecimientos tengamos. El tratamiento de esta segunda pregunta se contempla al final del capítulo, una vez determinado el modelo con el que vamos a trabajar.

### **5.1.- Variable a explicar:**

La variable que deseamos explicar es la cuota en tienda, esto es, la cuota de mercado en los establecimientos donde está presente la marca.

Los datos utilizados han sido proporcionados por Nielsen; son datos del mercado español de cacao soluble. Nielsen ha borrado cualquier referencia a las marcas que operan en el mercado para garantizar el anonimato de sus clientes.

¿Por qué la cuota en tienda y no la cuota de mercado? La razón es que cuando se trabaja con datos procedentes de un panel de detallistas, la evolución de un mercado se mide básicamente por dos variables:

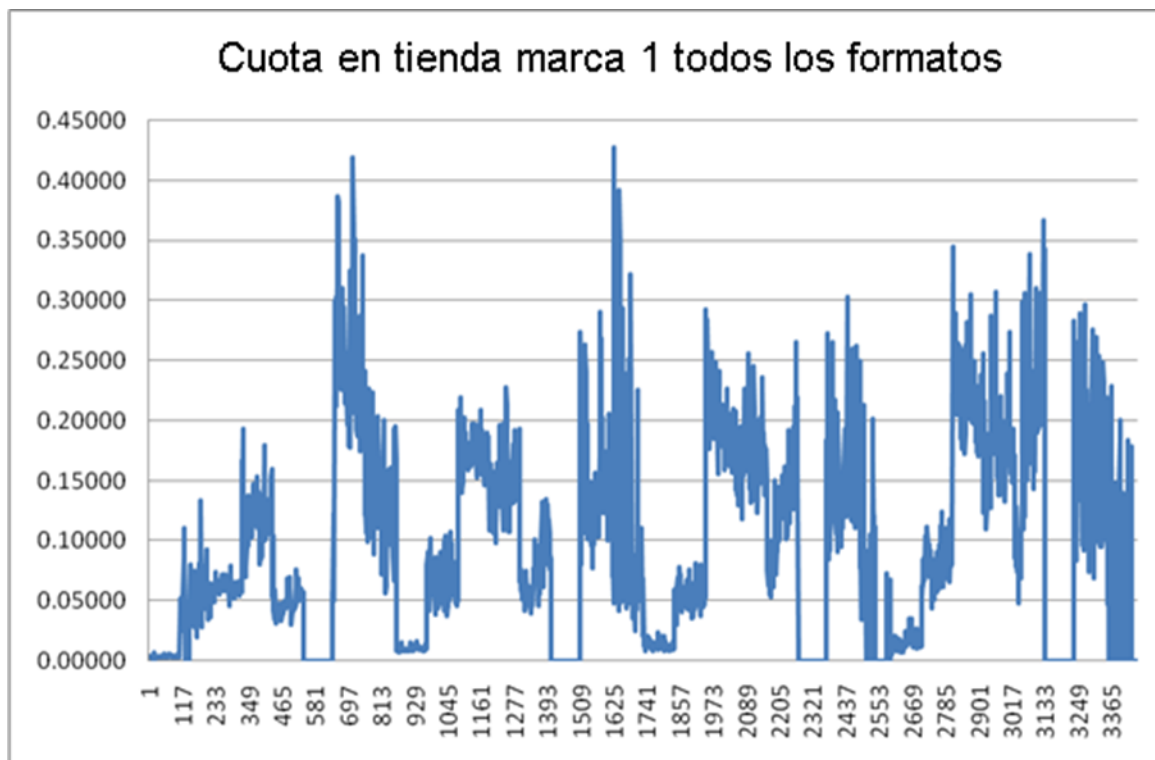
- La distribución, número de establecimientos que trabajan el producto y
- La cuota en tienda, cuota de mercado en los puntos de venta donde está presente la marca.

Esto es, un mercado crece porque hay más establecimientos que venden el producto, o porque los establecimientos que venden el producto, venden más, o bien es una conjunción de los dos factores. Trabajar con la cuota de mercado directamente supone añadir una fuente de heterogeneidad más a los datos; la distribución, por esta razón es preferible trabajar con una variable más homogénea y que proporcione información sobre los que está sucediendo en los puntos de venta. La forma de calcularla es:

$$\text{Cuota en tienda marca } i \text{ formato } j = \frac{\text{Ventas en tienda marca } i \text{ formato } j}{\text{Ventas Totales de los establecimientos donde está presente la marca } i \text{ formato } j}$$

Además la cuota de mercado es una entelequia. Es decir, cuando actuamos como compradores y realizamos una compra, esa compra se materializa en un producto concreto con un formato concreto y en un establecimiento concreto; posteriormente cuando queremos analizar el mercado hablamos de cuota de mercado como la suma de las ventas de una marca dividida por la suma de las ventas de todas las marcas. Pero realmente las compras se han materializado en productos, formatos y establecimientos concretos. En este sentido trabajar con la cuota de mercado supone trabajar con una variable más heterogénea que la cuota en tienda.

El gráfico siguiente muestra la evolución de la cuota en tienda de la marca 1 a lo largo de los diferentes periodos (semanas) y teniendo en cuenta los diferentes formatos de presentación del producto.



El gráfico proporciona la evidencia de un problema que es preciso resolver; la variable Cuota en tienda toma valor 0; esto es, no se ha producido ninguna venta. La razón es que la marca o bien no trabaja el formato considerado y es incorporado más tarde (tamaño muestral más pequeño) o bien el formato aparece y desaparece del establecimiento (menor tamaño muestral nuevamente). La solución adoptada es eliminar del análisis aquellos formatos que no tengan un mínimo de 108 observaciones. De esta manera se logra que el panel esté equilibrado, los formatos a estudiar tienen todos la misma muestral temporal.

Las consecuencias de esta decisión son:

- Se elimina del análisis el formato de 1500 a 2999 gr. de todos los establecimientos. Este formato es lanzado al mercado en la observación 100. Por lo tanto, sólo tendríamos 8 observaciones para analizarlo.
- Se elimina del análisis el formato de más de 5000 gr. en los establecimientos Supermercados de 400 a 999 m. y en los Supermercados de 100 a 399 m. El formato tiene un comportamiento intermitente, aparece y desaparece en el



canal, quizá se deba a que es un formato muy grande y los establecimientos pequeños no son los más adecuados para su venta.

- Se elimina del análisis el formato de más de 5000 gr. en los Hipermercados y en los Supermercados de 1000 a 2499 m. En este caso la razón es que la marca a estudiar es la única marca presente; no tiene competencia entre el grupo de Competidor Principal.



En consecuencia nos encontramos con que la variable a explicar: Cuota en tienda de la marca *1* va a ser tratada como una variable de panel con los siguientes elementos:

- Los elementos del panel van a estar representados por la letra *i* que indica los formatos de la marca a analizar en los distintos tipos de establecimientos. Así los datos serán:

<i>i</i>	ESTABLECIMIENTO	FORMATO
1	HIPER	Hasta 350
2	HIPER	De 350 a 499
3	HIPER	De 500 a 899
4	HIPER	De 900 a 1199
5	HIPER	De 1200 a 1499
6	HIPER	De 3000 a 4999
7	SUPERMERCADO DE 1000 A 2499 m	Hasta 350
8	SUPERMERCADO DE 1000 A 2499 m	De 350 a 499
9	SUPERMERCADO DE 1000 A 2499 m	De 500 a 899
10	SUPERMERCADO DE 1000 A 2499 m	De 900 a 1199
11	SUPERMERCADO DE 1000 A 2499 m	De 1200 a 1499
12	SUPERMERCADO DE 1000 A 2499 m	De 3000 a 4999
13	SUPERMERCADO DE 400 A 999 m	Hasta 350
14	SUPERMERCADO DE 400 A 999 m	De 350 a 499
15	SUPERMERCADO DE 400 A 999 m	De 500 a 899
16	SUPERMERCADO DE 400 A 999 m	De 900 a 1199
17	SUPERMERCADO DE 400 A 999 m	De 1200 a 1499
18	SUPERMERCADO DE 400 A 999 m	De 3000 a 4999
19	SUPERMERCADO DE 100 A 399 m	Hasta 350
20	SUPERMERCADO DE 100 A 399 m	De 350 a 499
21	SUPERMERCADO DE 100 A 399 m	De 500 a 899
22	SUPERMERCADO DE 100 A 399 m	De 900 a 1199
23	SUPERMERCADO DE 100 A 399 m	De 1200 a 1499
24	SUPERMERCADO DE 100 A 399 m	De 3000 a 4999

- Dentro de cada elemento encontramos 108 observaciones ordenadas en el tiempo con una periodicidad semanal. La primera observación corresponde a la semana 17.

## 5.2.- Factores explicativos.

Antes de enunciar las variables a utilizar como explicativas en el análisis, es preciso realizar una puntualización como consecuencia de la técnica que pretendemos usar. La estimación por datos de panel exige que las X estén presentes en todos los individuos y a lo largo de todo el periodo considerado.

En nuestro caso supone que todas las marcas con todos sus formatos estén presentes en todos los establecimientos y a lo largo de todos los periodos. Nada

más lejos de la realidad, en la información utilizada nos encontramos con casos donde una marca está presente en un tipo de establecimiento pero no en otro, y si está presente en un formato no significa que deba estarlo en todos los formatos considerados. Para resolver esta situación pensamos en definir a las variables explicativas respecto al **PRINCIPAL COMPETIDOR**. Esto significa que en unos establecimientos o formatos el principal competidor puede ser la marca 4 y en otros puede serlo la marca 5 y sin embargo sus acciones aparecerán de forma conjunta en el análisis, como si se tratara de una sola marca.

Esta decisión que si bien posibilita la utilización de la metodología de datos de panel, producirá un modelo más simplista y más alejado de la realidad en los establecimientos más grandes donde el nivel de competencia es más elevado, mientras que en los establecimientos pequeños estará más próxima a la realidad.

De acuerdo a esta decisión, eliminamos del análisis las marcas 5 y 6 ya que no están presentes en todos los tipos de establecimientos o bien donde están no constituyen el principal competidor. Por último el COMPETIDOR PRINCIPAL estará constituido por las marcas 4, 3 y 2; el núcleo más importante está constituido por la marca 4 excepto en el formato de hasta 350 gr. en el que la marca 3 tiene una mayor importancia en todos los tipos de establecimientos. La marca 2 la utilizamos como COMPETIDOR PRINCIPAL en 7 periodos en el formato de 900 a 1199 gr. en Hipermercados, la razón es que se trata de la principal competidora en este formato en esos periodos, posteriormente, la marca 4 lanzó el formato de 900 a 1199 gr. y se convirtió en el competidor principal de la marca 1, que es la que queremos estudiar.

Los dos factores a utilizar como explicativos en el análisis son RESTO DE FABRICANTES Y MARCAS DEL DISTRIBUIDOR, ambos factores tienen una característica común con el factor COMPETIDOR PRINCIPAL, son fruto de la agregación de distintas marcas. Así el caso de Resto de Fabricantes está constituido por pequeños fabricantes cuya importancia individual es pequeña y son agrupados en una rúbrica más importante, donde ya su importancia conjunta no es desdeñable.

El caso de las MARCAS DEL DISTRIBUIDOR recoge aquellos productos comercializados por las distintas cadenas con su marca privada. Como no existe una

única cadena de establecimientos, nos encontramos pues con distintas marcas del distribuidor; por lo tanto, de nuevo, la MARCA DEL DISTRIBUIDOR constituye una agregación de distintas marcas de distribuidores.

Como conclusión el objetivo perseguido es explicar la Cuota en tienda de la marca 1 en todos sus formatos, en un entorno competitivo formado por el PRINCIPAL COMPETIDOR, RESTO DE FABRICANTES Y MARCAS DEL DISTRIBUIDOR. Es sobre este entorno donde definiremos las variables a utilizar como explicativas.

### 5.3.- Variables explicativas.

- ***Diferencial de precio respecto al principal competidor:***  $x_2$

Con esta variable medimos la relación de precios existente entre el competidor principal y la marca 1. La definimos como:

$$x_2 = \frac{\text{Precio Principal Competidor}}{\text{Precio Marca 1}}$$

Cuando toma valores menores que 1, quiere decir que el Principal Competidor vende a un precio inferior que la marca 1. Cuando toma valores superiores a 1 se da la situación contraria; el Principal Competidor vende a un precio superior a la marca 1.

- ***Diferencial de precios respecto al Resto de Fabricantes.***  $x_3$

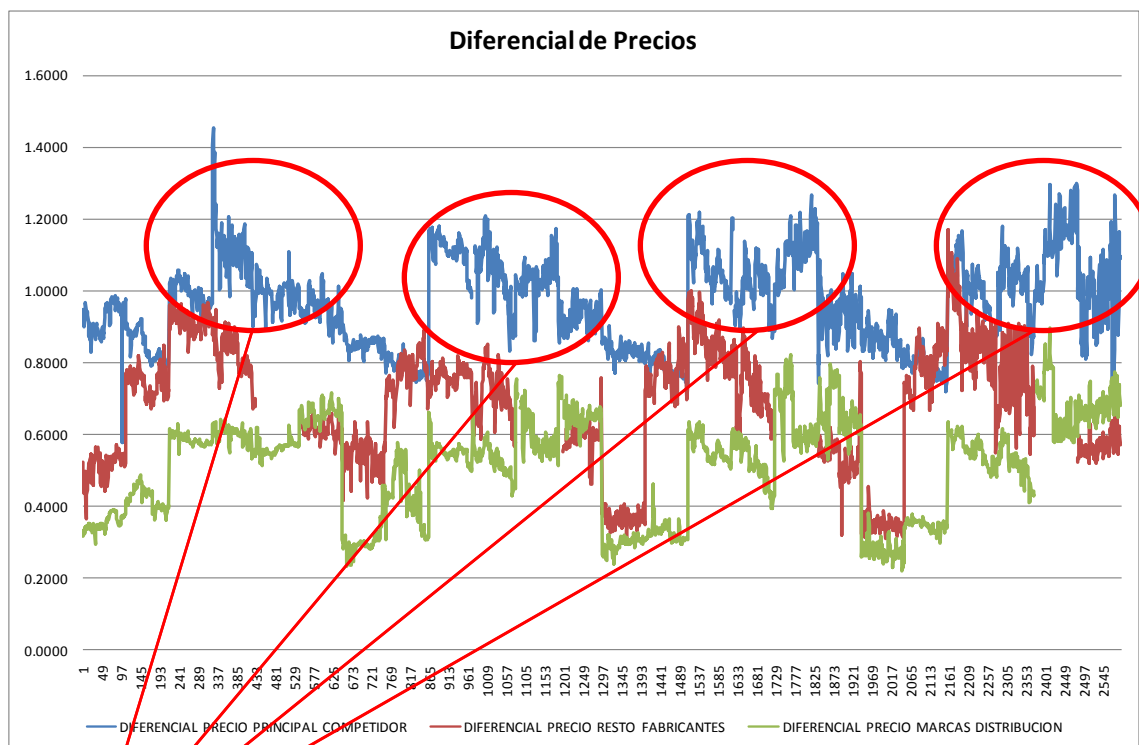
$$x_3 = \frac{\text{Precio Resto Fabricantes}}{\text{Precio Marca 1}}$$

- ***Diferencial de precios respecto a la Marca del Distribuidor:***  $x_4$

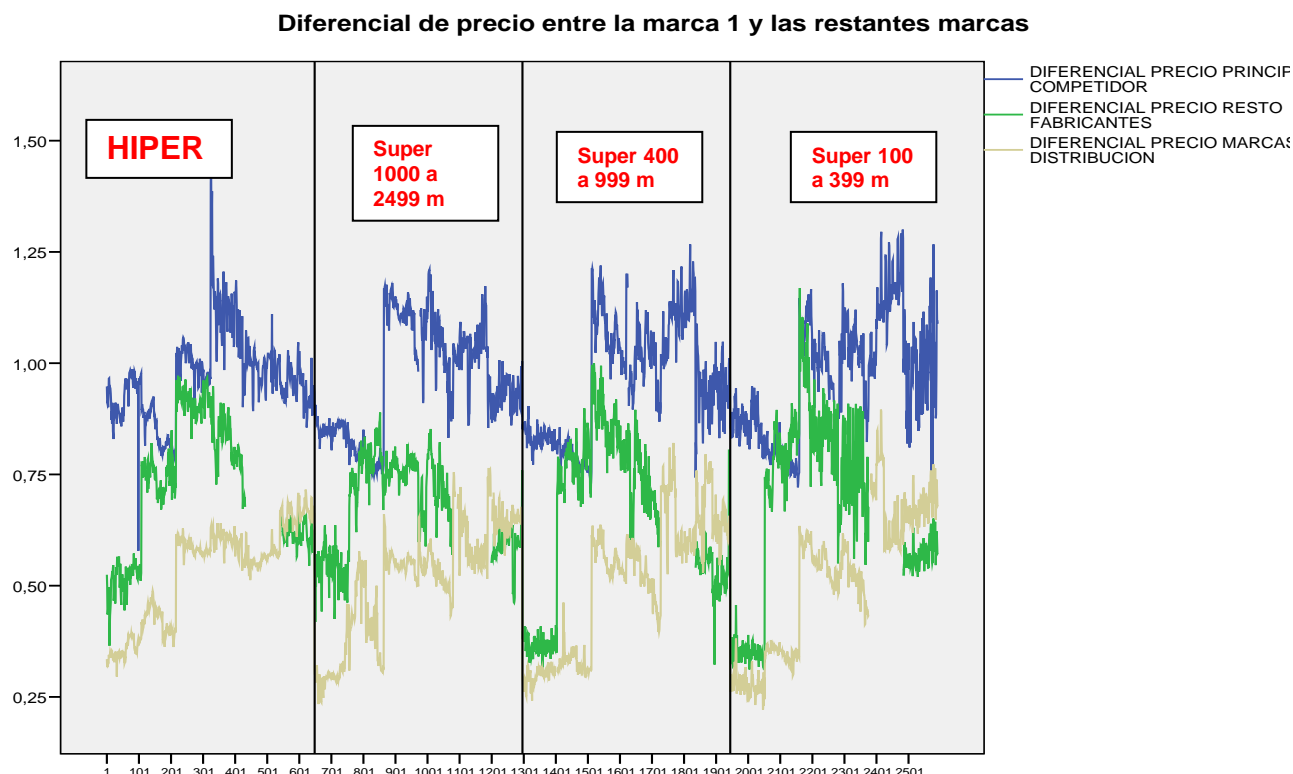
$$x_4 = \frac{\text{Precio Marcas Distribución}}{\text{Precio Marca 1}}$$

La interpretación de estas dos variables es exactamente la misma a primera.

En principio los Parámetros  $\beta$  asociados a estas variables deberían ser positivos. Valores inferiores a 1 supondrían que la marca 1 es más cara que sus competidores y por lo tanto debería tener una cuota en tienda inferior y al contrario, valores superiores a 1 significa que la marca 1 es más barata que su competencia y debería tener una cuota en tienda superior. Por tanto la relación que existe entre la variable  $y$  y las variables  $x_1, x_2, x_3$  es directa y nos llevaría a  $\beta$  con signos positivos.



Situaciones donde el Principal Competidor supera en precio a la Marca 1



El gráfico nos muestra como el diferencial de precios entre la marca 1 y sus competidores mantiene un comportamiento bastante similar independientemente del tipo de establecimiento. Esta semejanza es más acusada cuando sólo comparamos los supermercados.

- **Distribución Ponderada Marca 1:**  $x_1$

Es un índice que nos informa sobre la distribución (% de establecimientos que trabajan el producto) y la calidad de esta distribución (importancia que tienen en las ventas los establecimientos donde está presente la marca). La forma de calcularlo es:

$$x_1 = \frac{\text{Ventas en valor de la marca}}{\text{Ventas totales en valor de los establecimientos donde está la marca}}$$

El parámetro  $\beta$  asociado debería tener signo positivo; esto es, incrementos en la variable  $x_4$  deberían suponer un aumento del mismo signo en la variable explicativa.

- **Distribución Ponderada Principal Competidor:**  $x_5$

La interpretación de esta variable es la misma que en el caso anterior. Respecto al signo del parámetro asociado no tiene una interpretación clara; así podría tener signo negativo; en el sentido de que un aumento en la Distribución Ponderada del Principal Competidor debería producir una reducción de la cuota en tienda de la marca 1. Pero también podría tener signo positivo; de tal manera que el efecto llamada que se produce en las tiendas importantes, hace que todas las marcas vendan más. Por esta razón no podemos establecer un signo a priori para esta variable.

- **Distribución Ponderada Resto Fabricantes:**  $x_6$

- **Distribución Ponderada Marcas de la Distribución:**  $x_7$

- **Intensidad de la promoción de Precio de la Marca 1:**  $x_8$

$$x_8 = \frac{\text{Precio en promoción marca 1}}{\text{Precio habitual marca 1}}$$

Con esta variable estamos midiendo la intensidad de la promoción de precios realizada; esto es, que magnitud tiene la reducción de precios realizada respecto al precio habitual de la marca 1.

El valor máximo que tomará será 1 cuando coincidan el precio en promoción y el precio habitual; en este caso NO se está realizando ninguna promoción en precio. Los restantes valores serán inferiores a 1, siendo más pequeños en la medida de que la promoción de precios sea más intensa.

El signo del parámetro asociado a esta variable debería ser negativo, ya que valores bajos de la variable, gran intensidad promocional, deberían asociarse a niveles elevados de la cuota en tienda de la marca 1.

- **Intensidad de la promoción de Precio del PRINCIPAL COMPETIDOR:**  $x_{15}$

Esta variable mide la intensidad de la promoción en precios del Competidor Principal. Su rango de variación es el mismo que la variable  $x_8$  ahora bien, el signo asociado a su parámetro debería positivo, totalmente el contrario al de la variable  $x_8$ , esto es, valores bajos de la variable (gran intensidad en la promoción de precios) deberían asociarse a valores bajos de la cuota en tienda. Es entendible que acciones de precio de la competencia produzcan disminuciones en la cuota de mercado de la marca objeto de estudio.

- **Intensidad de la promoción de Precio del RESTO DE FABRICANTES:**  $x_{22}$

La interpretación y el signo asociado a esta variable es el mismo que el de la variable  $x_{15}$

- **Intensidad de la promoción de Precio de las MARCAS de DISTRIBUCION:**  
 $x_{29}$

La interpretación y el signo asociado a esta variable es el mismo que el de las variables  $x_{15}$  y  $x_{22}$

- **Cobertura de la Promoción Folleto y Exposición Especial de la Marca 1:**  
 $x_9$

La promoción Folleto y Exposición Especial es una promoción conjunta, en el que la marca recibe un tratamiento especial en el punto de venta, tanto desde el punto de vista de su ubicación -un espacio y lugar privilegiado para aproximarla al consumidor- como desde el punto de vista de la comunicación del establecimiento; todos los hogares cercanos al punto de venta recibe un folleto con las ofertas, que en un determinado momento, hay en el establecimiento.



La cobertura hace referencia al porcentaje de las ventas totales realizadas con este tipo de promoción.

El signo esperado a esta variable debería ser positivo; esto es, coberturas elevadas en este tipo de promoción deberían estar vinculadas a cuotas de mercado elevadas.

- **Cobertura de la Promoción sólo Exposición Especial de la Marca 1:**  $x_{10}$

Esta variable recoge idéntica información a la variable  $x_9$  pero haciendo referencia exclusivamente a la Exposición Especial. El signo esperado debe ser positivo.

- **Cobertura de la Promoción sólo Folleto de la Marca 1:**  $x_{11}$

En este caso, la variable está recogiendo la información relativa a la intensidad de la promoción sólo folleto realizada por la marca 1 en aquellos establecimientos donde está presente.

El signo asociado a esta variable de ser positivo, indicando una relación directa entre la Cuota en tienda y la variable.

- **Cobertura Otras Promociones Marca 1:**  $x_{12}$

Bajo la denominación de Otras Promociones se incluyen el conjunto de acciones realizadas por la Marca 1 que no han sido recogidas explícitamente en los tipos anteriores de promociones. Recoge acciones como las promociones de sorteo, de regalo diferido, etc. que no aparecen de forma explícita en los datos actuales.

En este caso, la variable nos proporciona información sobre la intensidad con la que se han utilizado estas acciones por la marca 1. Es esperable que el parámetro asociado a esta variable tenga signo positivo.

- **Cobertura Reducción Temporal de Precio Marca 1**  $x_{13}$

La reducción temporal de precio se refiere a la típica promoción realizada por las marcas, esto es, en un momento determinado y en un punto de venta concreto nos encontramos que la marca tiene un precio inferior al habitual.

Con la variable  $x_{13}$  se recoge esta información relativa a la marca 1, y más concretamente el porcentaje de establecimientos que realizan este tipo de promoción.

El signo del parámetro asociado a esta variable debe ser positivo. Cuantos más establecimientos realicen este tipo de acción mayor deberá ser la cuota de mercado de la marca 1.

- **Cobertura Regalo Directo Marca 1:**  $x_{14}$

La promoción Regalo Directo hace referencia a la forma de entregar el regalo en el que está basada la promoción. Cuando se habla de regalo directo significa que el consumidor se lleva el regalo al comprar el producto. El regalo puede ir en el interior del producto, a recoger en la caja del punto de venta o cualquier otra fórmula.

Esta promoción suele tener una efectividad mayor que otras promociones como los sorteos o las promociones de regalo diferido, aunque su coste y complicación logística es sensiblemente superior.

La variable  $x_{14}$  recoge esta información relativa a la marca 1. Es esperable que el signo asociado sea positivo.

El conjunto de variable que viene a continuación recogen el mismo tipo de información que las anteriores pero referidas al Principal Competidor, al Resto de Fabricantes y a Marcas de la Distribución; en una palabra, la competencia.

Los signos de los parámetros asociados a estas variables deberían ser NEGATIVOS, estableciendo una relación inversa entre la cuota de mercado de la marca 1 y las acciones de marketing que realiza la competencia.

- **Cobertura Promoción Folleto y Exposición Especial Principal Competidor:**  $x_{16}$
- **Cobertura Promoción Solo Exposición Especial Principal Competidor:**  
 $x_{17}$
- **Cobertura Promoción Solo Folleto Principal Competidor:**  $x_{18}$
- **Cobertura Otras Promociones Principal Competidor:**  $x_{19}$
- **Cobertura Reducción Temporal De Precio Principal Competidor:**  $x_{20}$
- **Cobertura Regalo Directo Principal Competidor:**  $x_{21}$
- **Cobertura Promoción Folleto Y Exposición Especial Resto De Fabricantes:**  $x_{23}$
- **Cobertura Promoción Solo Exposición Especial Resto De Fabricantes:**  
 $x_{24}$
- **Cobertura Promoción Solo Folleto Resto De Fabricantes:**  $x_{25}$
- **Cobertura Otras Promociones Resto De Fabricantes:**  $x_{26}$
- **Cobertura Reducción Temporal De Precios Resto De Fabricantes:**  $x_{27}$
- **Cobertura Regalo Directo Resto De Fabricantes:**  $x_{28}$
- **Cobertura Promoción Folleto Y Exposición Especial Marcas Distribución:**  $x_{30}$
- **Cobertura Promoción Solo Exposición Especial Marcas Distribución:**  $x_{31}$
- **Cobertura Promoción Solo Folleto Marcas Distribución:**  $x_{32}$
- **Cobertura Otras Promociones Marcas Distribución:**  $x_{33}$
- **Cobertura Reducción Temporal De Precios Marcas Distribución:**  $x_{34}$
- **Cobertura Regalo Directo Marcas Distribución:**  $x_{35}$

#### **5.4.- ESTIMACIÓN DEL MODELO.**

Una vez determinadas la variable a explicar y las posibles variables explicativas, parece obvio que los pasos siguientes vayan encaminados hacia la estimación del modelo. La dificultad en nuestro caso es que nos encontramos con posibles modelos a estimar y diferentes métodos a utilizar. En este capítulo iremos trazando el camino para llegar a un único modelo y con un método de estimación determinado, argumentando porque se van desechando las otras opciones.

En un primer momento estaremos el modelo de forma ESTÁTICA, todas las variables están referidas al mismo momento del tiempo, para luego abrirnos a un tratamiento dinámico de los datos.

## 5.4.1- Estimación estática del modelo

### 5.4.1.1.- Estadísticos descriptivos

. summarize

Variable	Obs	Mean	Std. Dev.	Min	Max
id	2592	1296.5	748.3903	1	2592
describ	0				
panel	2592	12.5	6.923522	1	24
year	2592	2005.852	.7048137	2005	2007
week	2592	26.2037	14.80949	1	52
y	2592	.1147342	.0796766	.0025886	.4195191
x1	2592	59.84684	31.96321	1	100
x2	2592	.9639289	.1210044	.5784526	1.453192
x3	2592	.7351587	.1942195	.3116308	1.168932
x4	2592	.5180124	.144244	.2216561	1
x5	2592	65.2365	33.36234	0	100
x6	2592	43.22878	37.34634	0	100
x7	2592	38.57909	30.8201	0	94
x8	2592	.9573121	.0434296	.6884892	1
x9	2592	1.561888	3.641316	0	43.43435
x10	2592	3.210182	6.277277	0	48
x11	2592	3.241711	6.822419	0	65.21739
x12	2592	8.876263	14.09254	0	86.36364
x13	2592	9.737739	10.8391	0	70.14925
x14	2592	6.699429	16.45628	0	89.79592
x15	2592	.9641888	.0414366	.535	1
x16	2592	1.825011	5.730573	0	100
x17	2592	2.534777	6.165083	0	66.66666
x18	2592	3.891771	9.146211	0	100
x19	2592	2.163112	7.582771	0	72
x20	2592	9.742096	11.87358	0	133.3333
x21	2592	5.066153	10.81839	0	116.6667
x22	2592	.9220193	.1359206	.0444874	1
x23	2592	.4190274	2.164954	0	50
x24	2592	1.561304	6.935371	0	100
x25	2592	1.409191	4.634495	0	55.55556
x26	2592	4.804803	14.75811	0	100
x27	2592	6.736501	10.03241	0	100
x28	2592	.1905036	1.334559	0	33.33333
x29	2592	.9721828	.0557611	.4249292	1
x30	2592	.5556686	2.941241	0	100
x31	2592	1.973831	6.506503	0	100
x32	2592	3.519566	10.56723	0	100
x33	2592	.0494263	.8578905	0	22.95082
x34	2592	7.956456	17.21356	0	100
x35	2592	.0044387	.0953232	0	2.352941
time	2592	2409.5	31.18159	2356	2463

Un análisis más pormenorizado de la variable dependiente arroja los siguientes datos:

cuota en tienda marca					
Percentiles	Smallest				
1%	.0033445	.0025886			
5%	.0085061	.0025928			
10%	.0113381	.0026252	Obs	2592	
25%	.0514506	.0026385	Sum of wgt.	2592	
50%	.1063456		Mean	.1147342	
		Largest	Std. Dev.	.0796766	
75%	.1759966	.3822314			
90%	.2199801	.3866349	Variance	.0063484	
95%	.2508859	.4123541	Skewness	.4644941	
99%	.3091243	.4195191	Kurtosis	2.512051	

Especialmente interesantes resultan los datos relativos a la asimetría (cercano a 0) y curtosis (cercano a 3) que nos hacen concebir esperanzas respecto a la normalidad de la variable dependiente.

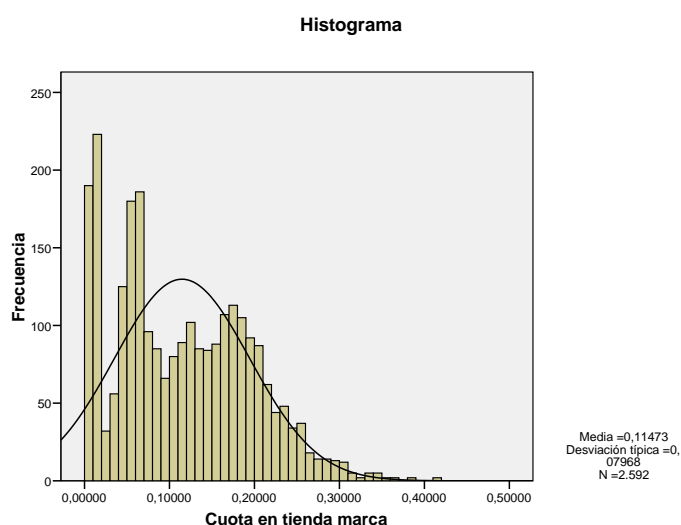
Para resolver esta cuestión se plantea un test de Kolmogorov Smirnov, en la versión recogida por Siegel (1994, 69) sobre la variable **y** dando como resultado que debemos rechazar la hipótesis de normalidad planteada sobre esta variable (cuota en tienda)

### Prueba de Kolmogorov-Smirnov para una muestra

		Cuota en tienda marca
N		2592
	Media	.11473
Parámetros normales(a,b)		42
	Desviación típica	.07967
Diferencias más extremas	Absoluta	.096
	Positiva	.096
	Negativa	-.080
Z de Kolmogorov-Smirnov		4.892
Sig. asintót. (bilateral)		.000

a La distribución de contraste es la Normal.

b Se han calculado a partir de los datos.



### 5.4.1.2.- Correlaciones entre las variables

	x1	x2	x3	x4	x5	x6	x7
x1	1.0000						
x2	0.5675	1.0000					
x3	0.1717	0.4838	1.0000				
x4	0.4405	0.6144	0.4494	1.0000			
x5	0.0030	-0.4543	-0.0837	-0.4633	1.0000		
x6	-0.0623	-0.2879	-0.1289	-0.5191	0.7138	1.0000	
x7	0.4745	0.2670	0.3582	0.1650	0.3165	0.2447	1.0000
x8	-0.3323	-0.3032	-0.2265	-0.1450	0.0714	-0.0425	-0.2782
x9	0.2232	0.2908	0.0894	0.2261	-0.1263	-0.1242	0.2513
x10	0.3005	0.1631	-0.0908	0.3534	-0.0685	-0.2875	0.1973
x11	0.2517	0.3913	0.1421	0.2601	-0.1320	-0.1564	0.1925
x12	0.4844	0.4554	0.1757	0.2740	-0.0464	0.1551	0.5314
x13	0.2699	0.4507	0.1328	0.3194	-0.2653	-0.2274	0.1422
x14	0.1382	-0.0439	-0.2652	0.3983	-0.1947	-0.4052	-0.1133
x15	0.0394	-0.0231	-0.1670	-0.0766	-0.0068	-0.0137	-0.1155
x16	0.0329	-0.1213	-0.0461	0.0766	-0.0236	-0.0604	0.0809
x17	0.1194	-0.0923	-0.0152	0.1451	-0.0132	-0.1380	0.0399
x18	0.0791	-0.1260	0.0022	0.0785	0.0833	0.0063	0.1428
x19	0.1437	-0.1164	-0.2670	-0.2844	0.2785	0.3270	-0.1648
x20	0.0318	-0.1560	0.0046	0.0399	0.0767	0.0172	0.0886
x21	0.0185	-0.1076	-0.1151	0.0452	0.0772	-0.0731	0.0022
x22	0.1216	0.3277	0.3741	0.5118	-0.3274	-0.4055	0.1539
x23	0.1191	0.0132	0.0808	0.0384	0.1543	0.1482	0.2224
x24	0.0288	-0.0391	-0.1088	0.0404	0.0645	-0.0001	0.0651
x25	-0.0468	-0.0695	0.0488	-0.1081	0.1611	0.2005	0.1684
x26	0.1628	-0.0654	-0.0918	0.1138	0.2561	-0.0548	0.2322
x27	-0.1012	-0.1793	-0.1260	-0.2634	0.3376	0.4303	0.1168
x28	-0.0250	-0.0935	-0.0014	-0.0190	0.1326	0.0848	0.0701
x29	0.1123	0.1305	-0.0142	0.0569	-0.1474	-0.1674	-0.1467
x30	0.0894	0.0860	0.0340	0.0611	-0.0334	-0.0001	0.1108
x31	0.0668	0.1186	0.0638	0.0544	-0.0426	0.0262	0.0718
x32	0.1772	0.1511	0.0559	0.1222	-0.0198	-0.0525	0.2122
x33	-0.1061	-0.0483	0.0012	-0.0282	0.0595	0.0860	0.0271
x34	0.0261	0.0086	-0.0557	-0.0582	-0.0386	-0.0153	0.0033
x35	0.0568	0.0212	-0.0087	0.0316	0.0405	0.0074	0.0731

	x8	x9	x10	x11	x12	x13	x14
x8	1.0000						
x9	-0.0288	1.0000					
x10	0.0944	0.1881	1.0000				
x11	0.0322	0.3575	0.1895	1.0000			
x12	-0.3261	0.1067	0.0617	0.0868	1.0000		
x13	-0.0925	0.4176	0.2702	0.6002	0.1205	1.0000	
x14	0.2419	0.0577	0.6141	0.0813	-0.1206	0.1384	1.0000
x15	0.0210	-0.0084	0.0503	-0.0708	-0.0659	-0.0510	0.0215
x16	0.0186	0.0392	0.1253	-0.0017	0.1028	0.0151	0.1946
x17	0.0628	0.0111	0.2252	0.0284	0.0488	0.0079	0.2226
x18	-0.0109	0.0303	0.1165	0.0125	0.0719	0.0394	0.1828
x19	0.0654	-0.0628	-0.0731	-0.0896	-0.1387	-0.1148	-0.0599
x20	-0.0177	0.0472	0.1128	0.0340	0.0224	0.0563	0.1245
x21	0.0707	0.0171	0.2364	0.0105	-0.0426	-0.0252	0.1958
x22	-0.1064	0.1464	0.2048	0.1357	0.1462	0.1984	0.2259
x23	-0.1117	-0.0089	0.0276	-0.0092	0.1724	-0.0156	-0.0171
x24	0.0750	0.0229	0.1682	0.0340	0.0581	0.0057	0.1067
x25	-0.0576	-0.0361	-0.0509	-0.0391	0.0862	-0.0419	-0.1043
x26	0.1266	0.1879	0.4756	0.0922	0.0394	-0.0024	0.3014
x27	-0.0291	-0.0856	-0.0975	-0.1060	0.0786	-0.0671	-0.1689
x28	0.0513	-0.0155	0.0728	-0.0216	-0.0271	-0.0358	0.1210
x29	-0.0338	0.0131	0.0073	0.0229	-0.0337	0.0202	-0.0000
x30	-0.0682	0.0517	0.0233	0.0511	0.1586	0.0618	-0.0002
x31	-0.0410	0.0261	0.0313	0.0602	0.1011	-0.0004	0.0013
x32	-0.0522	0.0291	0.1782	0.1069	0.1714	0.1076	0.1216
x33	0.0567	-0.0247	-0.0295	-0.0274	-0.0363	-0.0518	-0.0235
x34	-0.0070	0.0022	0.0155	0.0171	0.0490	0.0657	0.0205
x35	0.0138	0.0073	0.0992	0.1074	0.0305	0.0213	0.0383

	x15	x16	x17	x18	x19	x20	x21
x15	1.0000						
x16	-0.0661	1.0000					
x17	-0.0368	0.0662	1.0000				
x18	-0.0649	0.2539	0.0823	1.0000			
x19	0.1122	-0.0790	-0.0629	-0.0848	1.0000		
x20	-0.0859	0.3946	0.2208	0.5756	-0.0970	1.0000	
x21	0.1674	0.1052	0.1395	0.0822	-0.0465	0.0808	1.0000
x22	-0.1220	0.0790	0.1287	0.0840	-0.3880	0.0848	0.0201
x23	0.0281	0.0421	0.0037	0.0611	-0.0315	0.0346	0.1274
x24	0.0074	0.0228	0.0526	0.0368	0.0114	0.0346	0.1552
x25	-0.0378	-0.0143	-0.0495	0.0215	-0.0281	-0.0174	-0.0185
x26	-0.0000	0.1516	0.1995	0.1198	0.0332	0.0999	0.1519
x27	-0.0107	-0.0073	-0.0352	0.0078	0.0863	0.0362	-0.0433
x28	-0.0409	0.0525	0.0406	0.0243	0.0514	0.0252	-0.0034
x29	0.0349	-0.0378	-0.0225	-0.0354	0.0698	-0.0046	-0.0760
x30	-0.0315	0.0119	0.1160	-0.0100	-0.0332	0.0580	-0.0245
x31	0.0126	-0.0294	0.0121	-0.0485	-0.0647	-0.0458	0.1144
x32	-0.0485	0.0503	0.0373	0.0979	-0.0710	0.0593	0.0328
x33	0.0191	-0.0154	-0.0162	-0.0235	-0.0164	-0.0040	0.0578
x34	0.0172	-0.0138	0.0196	-0.0129	-0.0129	-0.0037	0.0334
x35	-0.0122	0.0057	0.0068	0.0450	-0.0119	0.0263	0.0530

	x22	x23	x24	x25	x26	x27	x28
x22	1.0000						
x23	0.0294	1.0000					
x24	0.0254	0.0129	1.0000				
x25	-0.0285	0.2422	0.0030	1.0000			
x26	0.1032	-0.0188	0.1189	-0.0522	1.0000		
x27	-0.1197	0.1263	0.1717	0.3810	-0.0100	1.0000	
x28	0.0073	0.0363	0.1384	0.0191	0.0634	0.0517	1.0000
x29	-0.0157	-0.0572	-0.0124	-0.0396	-0.0124	-0.0378	-0.0014
x30	0.0390	-0.0089	0.0015	-0.0006	0.0249	-0.0101	-0.0184
x31	0.0402	-0.0113	0.0364	-0.0080	-0.0130	-0.0049	-0.0108
x32	0.0953	0.0075	0.0248	0.0064	0.0586	-0.0471	-0.0153
x33	-0.0086	-0.0064	0.0441	-0.0158	-0.0188	0.0223	0.0894
x34	-0.0172	-0.0167	-0.0200	0.0122	-0.0359	0.0261	-0.0277
x35	0.0266	0.1167	0.0091	-0.0142	0.0705	-0.0191	-0.0066

	x29	x30	x31	x32	x33	x34	x35
x29	1.0000						
x30	0.0182	1.0000					
x31	-0.0428	0.0317	1.0000				
x32	-0.0174	0.2211	-0.0008	1.0000			
x33	-0.0607	-0.0109	-0.0126	0.0055	1.0000		
x34	-0.2273	0.1226	0.0641	0.2680	0.0045	1.0000	
x35	0.0172	0.0269	-0.0022	0.0788	-0.0027	0.0041	1.0000

El estudio de la matriz de correlaciones entre las variables explicativas nos permite hacernos una idea de la multicolinealidad existente entre estas variables. Como puede apreciarse, el único coeficiente que podríamos decir que es elevado es el que comparten las variables x5 y x6 (0,7138). Por tanto podemos afirmar que no existe multicolinealidad entre las variables y ello, en gran medida, por la utilización de datos de panel que hace posible trabajar con muestras muy grandes sin tener que preocuparnos por la multicolinealidad.



### 5.4.1.3.- Variabilidad del modelo

Variable		Mean	Std. Dev.	Min	Max	Observations	
y	overall	.1147342	.0796766	.0025886	.4195191	N =	2592
	between		.0746453	.0038942	.2596048	n =	24
	within		.0317262	-.0160792	.3038207	T =	108
x1	overall	59.84684	31.96321	1	100	N =	2592
	between		31.86385	1.583333	99.33333	n =	24
	within		6.947696	24.75424	83.72647	T =	108
x2	overall	.9639289	.1210044	.5784526	1.453192	N =	2592
	between		.109562	.7851955	1.128401	n =	24
	within		.0559817	.6185551	1.305764	T =	108
x3	overall	.7351587	.1942195	.3116308	1.168932	N =	2592
	between		.1896081	.3499952	1	n =	24
	within		.0570493	.5155169	1.10459	T =	108
x4	overall	.5180124	.144244	.2216561	1	N =	2592
	between		.1339212	.771333	.6911187	n =	24
	within		.0601007	.385442	1.187969	T =	108
x5	overall	65.2365	33.36234	0	100	N =	2592
	between		32.92635	6.481481	99.62963	n =	24
	within		8.583312	-8.55054	97.20872	T =	108
x6	overall	43.22878	37.34634	0	100	N =	2592
	between		37.60918	0	99.47222	n =	24
	within		6.221797	8.663966	67.15471	T =	108
x7	overall	38.57909	30.8201	0	94	N =	2592
	between		30.44761	1.75	88.68519	n =	24
	within		7.817036	.8290895	72.2365	T =	108
x8	overall	.9573121	.0434296	.6884892	1	N =	2592
	between		.0254134	.8877662	.9984475	n =	24
	within		.0355943	.6863429	1.024112	T =	108
x9	overall	1.561888	3.641316	0	43.43435	N =	2592
	between		1.678338	0	5.30944	n =	24
	within		3.249414	-3.747552	39.68679	T =	108
x10	overall	3.210182	6.277277	0	48	N =	2592
	between		4.655987	0	21.72894	n =	24
	within		4.315232	-17.50866	29.72077	T =	108
x11	overall	3.241714	6.822419	0	65.21739	N =	2592
	between		2.928359	0	10.3347	n =	24
	within		6.190658	-7.092993	61.20614	T =	108
x12	overall	8.876263	14.09254	0	86.36364	N =	2592
	between		10.76765	0	29.02881	n =	24
	within		9.351235	-20.15255	88.69291	T =	108
x13	overall	9.737739	10.8391	0	70.14925	N =	2592
	between		5.156432	1.388889	18.00114	n =	24
	within		9.591423	-8.2634	67.3019	T =	108
x14	overall	6.699429	16.45628	0	89.79592	N =	2592
	between		14.21757	0	43.34427	n =	24
	within		8.775965	-35.64484	53.15109	T =	108
x15	overall	.9641888	.0414366	.535	1	N =	2592
	between		.0148183	.9375589	.9952523	n =	24
	within		.0388134	.5213875	1.02663	T =	108
x16	overall	1.825011	5.730573	0	100	N =	2592
	between		1.84991	0	7.195789	n =	24
	within		5.436785	-5.370778	98.97928	T =	108
x17	overall	2.534777	6.165083	0	66.66666	N =	2592
	between		2.510849	.1197573	9.948692	n =	24
	within		5.653692	-7.413915	66.26934	T =	108
x18	overall	3.891771	9.146211	0	100	N =	2592
	between		3.152068	0	10.64815	n =	24
	within		8.609758	-6.756377	93.24362	T =	108
x19	overall	2.163112	7.582771	0	72	N =	2592
	between		6.719981	0	30.30391	n =	24
	within		3.768973	-28.1408	43.8592	T =	108

x20	overall	9.742096	11.87358	0	133.3333	N =	2592
	between		3.252965	4.602432	16.82596	n =	24
	within		11.43841	-7.083862	127.1804	T =	108
x21	overall	5.066153	10.81839	0	116.6667	N =	2592
	between		3.884392	.0485009	15.48938	n =	24
	within		10.1278	-10.42323	113.5538	T =	108
x22	overall	.9220193	.1359206	.0444874	1	N =	2592
	between		.0953724	.6187276	1	n =	24
	within		.0987631	.1684814	1.303292	T =	108
x23	overall	.4190274	2.164954	0	50	N =	2592
	between		.9432393	0	4.533653	n =	24
	within		1.958077	-4.114626	48.95077	T =	108
x24	overall	1.561304	6.935371	0	100	N =	2592
	between		2.128017	0	8.812163	n =	24
	within		6.614978	-7.250858	99.70945	T =	108
x25	overall	1.409191	4.634495	0	55.55556	N =	2592
	between		1.561742	0	5.346263	n =	24
	within		4.374955	-3.937072	54.87925	T =	108
x26	overall	4.804803	14.75811	0	100	N =	2592
	between		12.0074	0	59.96869	n =	24
	within		8.920663	-55.16389	100.1752	T =	108
x27	overall	6.736501	10.03241	0	100	N =	2592
	between		5.018305	0	16.65092	n =	24
	within		8.746766	-8.872756	104.8846	T =	108
x28	overall	.1905038	1.334559	0	33.33333	N =	2592
	between		.3232647	0	1.19248	n =	24
	within		1.296481	-1.001977	32.33136	T =	108
x29	overall	.9721828	.0557611	.4249292	1	N =	2592
	between		.0204301	.9137626	.9999104	n =	24
	within		.0520495	.4457538	1.05842	T =	108
x30	overall	.5556686	2.941241	0	100	N =	2592
	between		.6356177	0	2.039382	n =	24
	within		2.874643	-1.483713	99.62974	T =	108
x31	overall	1.973831	6.506503	0	100	N =	2592
	between		2.026765	0	9.021202	n =	24
	within		6.196486	-7.047371	100.122	T =	108
x32	overall	3.519566	10.56723	0	100	N =	2592
	between		4.077739	0	18.13065	n =	24
	within		9.782013	-14.61108	100.3255	T =	108
x33	overall	.0494263	.8578905	0	22.95082	N =	2592
	between		.2421383	0	1.18623	n =	24
	within		8244795	-1.136804	21.81401	T =	108
x34	overall	7.956453	17.21356	0	100	N =	2592
	between		3.458949	2.156115	14.64356	n =	24
	within		16.8771	-6.687106	104.5614	T =	108
x35	overall	.0044387	.0953232	0	2.352941	N =	2592
	between		.0152789	0	.0624272	n =	24
	within		.0941419	-.0579885	2.313279	T =	108

Este cuadro permite observar la variabilidad de los datos de las distintas observaciones, distinguiendo entre las fuentes de dicha variabilidad. Así OVERALL refleja la variabilidad global de la variable, BETWEEN la variabilidad de la variable entre las distintas unidades del panel y WITHIN la variabilidad de la variable dentro de las unidades del panel. De tal manera que podemos agrupar las variables en dos grandes grupos:

- Variables que varían entre las distintos elementos del panel
- Variables cuya variabilidad se produce dentro de las unidades del panel.

Las variables con marco rojo, reflejan variabilidad, principalmente, ENTRE las unidades del panel. Y las variables enmarcadas en verde reflejan variabilidad DENTRO de las unidades del panel. Por lo tanto con el modelo se pretende explicar ambos tipos de variabilidad.

Ahora bien, la estimación del modelo pone de manifiesto un problema que es necesario abordar previamente. El problema es que si sustituimos las variables explicativas por sus posibles valores llegaríamos a obtener valores en la variable a explicar que están fuera de su rango. Esto es, la  $y$  podría tomar valores por debajo de 0 o por encima de 1; y esto, en la práctica, resulta imposible ya que la variable  $y$  representa una cuota de mercado por lo que su rango sólo puede oscilar en 0 y 1.

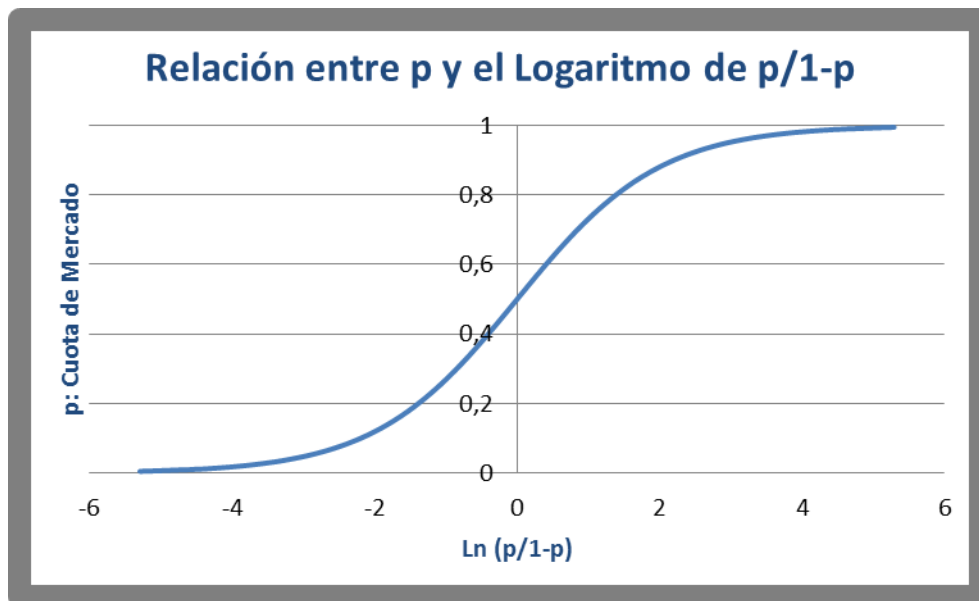
Una posible solución al problema puede ser restringir los valores  $y$  sólo observar aquellos que estén dentro del campo lógico de existencia. Este tipo de solución fue usada cuando se estimaba el comportamiento de una variable cualitativa mediante modelos lineales. Posteriormente, esta solución se desechó cuando empezaron a utilizarse los modelos logit.

La solución al problema planteado me la proporcionó el profesor Baltagi, en un curso que impartió en el mes de junio de 2010 en el Instituto de Estudios Fiscales en Madrid, y al que tuve la oportunidad de asistir. Baltagi me propuso resolver el problema de la siguiente forma: en lugar de modelizar la cuota de mercado ( $p$ )

modelizar la relación  $Ln\left(\frac{p}{1-p}\right)$ ; a partir de esta relación se puede tratar linealmente

su relación con las variables explicativas. Además, como ya se adelantaba en el primer capítulo, Leeflang y Wittink (2000) desarrollan una idea de Little respecto a la estructura de los modelos. Los modelos deben ser simples, completos respecto a las variables, adaptativos y robustos. La robustez puede lograrse con una estructura que restrinja las respuestas a un rango de valores que tenga sentido. Por tanto, si los valores de la variable explicativa están restringidos entre 0 y 1, como es el caso de la cuota de mercado, el modelo debería satisfacer estas mismas restricciones. Y podrían considerarse a este modelo como consistente. Es esta una razón más para incorporar la variable a explicar de acuerdo con la relación propuesta por Baltagi.

En el siguiente gráfico se muestra la relación existente entre  $p$  y  $\ln\left(\frac{p}{1-p}\right)$



Puede observarse que valores de  $p$  superiores al 50 % se corresponden con valores de  $\ln\left(\frac{p}{1-p}\right)$  positivos y valores de  $p$  inferiores al 50 % se corresponden con valores negativos. Por otro lado, resulta imposible obtener en  $p$  valores superiores a 1 e inferiores a 0, con lo que el problema enunciado queda resuelto.

Logproporcion				
Percentiles		Smallest		
1%	-5.6971	-5.954027		
5%	-4.758429	-5.952409		
10%	-4.468185	-5.939959	Obs	2592
25%	-2.914311	-5.934894	Sum of wgt.	2592
50%	-2.128625		Mean	-2.450834
		Largest	Std. Dev.	1.19269
75%	-1.54371	-.4800877		
90%	-1.265782	-.4614796	Variance	1.42251
95%	-1.093893	-.3542422	Skewness	-1.031416
99%	-.8042166	-.3247479	Kurtosis	3.334719

En esta tabla pueden verse los valores que toma la nueva variable a explicar. Para las variables explicativas nada ha cambiado por tanto los cuadros y gráficos anteriores continúan siendo válidos.

Veamos que sucede con la estimación de la nueva variable con los métodos que disponemos.

## 5.4.2- ESTIMACIÓN POR MCO AGRUPADOS.

El modelo a estimar es del tipo  $y_{it} = \alpha + \mathbf{x}'_{it}\beta + u_{it}$  que es modelo más restrictivo posible, establece que todos los coeficientes son iguales para los distintos individuos que componen el panel.

Si este modelo está correctamente especificado y los regresores incorrelacionados con el error, entonces puede estimarse CONSISTENTEMENTE por el método propuesto. Ahora bien si el modelo adecuado es el de Efectos Fijos, el estimador MCO agrupado resulta ser No consistente.

La estimación por MCO agrupados es la que se muestra a continuación.

Source	SS	df	MS			
Model	2946.62474	35	84.1892783			
Residual	739.097704	2556	.289161856			
Total	3685.72244	2591	1.42250963			

Number of obs =	2592
F( 35, 2556) =	291.15
Prob > F =	0.0000
R-squared =	0.7995
Adj R-squared =	0.7967
Root MSE =	.53774

logproporc~n	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
x1	-.0025274	.0005887	-4.29	0.000	-.0036818	-.0013731
x2	-.2432966	.1930804	-1.26	0.208	-.6219064	.1353132
x3	1.636362	.0909354	17.99	0.000	1.458048	1.814677
x4	.9255853	.1405039	6.59	0.000	.6500722	1.201098
x5	-.0066885	.0006594	-10.14	0.000	-.0079816	-.0053955
x6	.0010027	.0005487	1.83	0.068	-.0000732	.0020787
x7	.0054928	.0005899	9.31	0.000	.0043361	.0066496
x8	.3817528	.2876013	1.33	0.185	-.1822024	.945708
x9	.0379131	.003444	11.01	0.000	.0311599	.0446663
x10	.0266081	.0026018	10.23	0.000	.0215062	.03171
x11	.0150357	.0020847	7.21	0.000	.0109478	.0191236
x12	.0145386	.0011445	12.70	0.000	.0122944	.0167829
x13	.008858	.0013905	6.37	0.000	.0061313	.0115846
x14	.0146833	.0011131	13.19	0.000	.0125006	.016866
x15	.1188997	.2688868	0.44	0.658	-.4083584	.6461577
x16	.0084599	.0021228	3.99	0.000	.0042974	.0126224
x17	.0061703	.0019336	3.19	0.001	.0023786	.009962
x18	-.0056199	.0014738	-3.81	0.000	-.0085098	-.00273
x19	-.0439024	.0017865	-24.57	0.000	-.0474056	-.0403993
x20	-.0008473	.0012023	-0.70	0.481	-.0032049	.0015103
x21	-.0012597	.001082	-1.16	0.244	-.0033814	.0008619
x22	.7626727	.1002483	7.61	0.000	.5660965	.9592488
x23	-.0140135	.005317	-2.64	0.008	-.0244395	-.0035875
x24	.004096	.0016301	2.51	0.012	.0008996	.0072924
x25	.0065635	.0025745	2.55	0.011	.0015151	.0116119
x26	.0036847	.0009458	3.90	0.000	.0018302	.0055392
x27	-.0011681	.0013005	-0.90	0.369	-.0037183	.001382
x28	.0073334	.0082792	0.89	0.376	-.0089012	.0235681
x29	.1946226	.2082374	0.93	0.350	-.2137087	.6029538
x30	.0061371	.0037738	1.63	0.104	-.001263	.0135372
x31	.0094547	.0016825	5.62	0.000	.0061555	.0127539
x32	-.0008817	.001117	-0.79	0.430	-.003072	.0013086
x33	-.0135871	.0126311	-1.08	0.282	-.0383554	.0111812
x34	.0008247	.0006742	1.22	0.221	-.0004972	.0021467
x35	-.1873359	.1136026	-1.65	0.099	-.4100984	.0354266
_cons	-5.390621	.4744018	-11.36	0.000	-6.320872	-4.46037

A continuación vamos a comprobar si efectivamente el modelo planteado es adecuado; esto es si el modelo tiene un intercepto común y una pendiente común para todos los individuos.

Para ello vamos a utilizar un test de agrupabilidad propuesto por Hsiao (2002: 15 - 21) y al que hemos hecho referencia en el capítulo anterior. En este caso concreto el contraste que vamos a utilizar plantea como hipótesis nula que el modelo tiene un intercepto y una pendiente común a todos los individuos frente a la hipótesis alternativa de que los individuos no tienen intercepto común dando por sentado que la pendiente si es común. En términos paramétricos, las hipótesis a plantear son:

$$H_4 : \alpha_1 = \alpha_2 = \dots \alpha_N$$

$$\text{dado } \beta_1 = \beta_2 = \dots \beta_N$$

En términos econométricos, estamos testando si el modelo  $y_{it} = \alpha + \beta' x_{it} + u_{it}$  es el adecuado frente al modelo  $y_{it} = \alpha_i + \beta' x_{it} + u_{it}$

Este test está basado en los contraste de suma de residuos al cuadrado de la regresión lineal, muy comunes en econometría.

$$\text{El estadístico de contraste es: } F_4 = \frac{\frac{S_3 - S_2}{(N-1)}}{\frac{S_2}{N(T-1) - K}} \text{ donde}$$

- $S_3$  suma de los residuos al cuadrado de la regresión MCO agrupados  $y_{it} = \alpha + \beta' x_{it} + u_{it}$ .
- $S_2$  suma de los residuos al cuadrado del modelo  $y_{it} = \alpha_i + \beta' x_{it} + u_{it}$  que generalmente suele estimarse por Efectos Fijos.

Rechazaremos la hipótesis nula si  $F_4 \geq F_{(N-1), N(T-1)-K}^{(\alpha)}$

La estimación por efectos fijos aporta los siguientes resultados:

```
. xtreg logproporcion x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 x13 x14 x15 x16 x17 x
> 18 x19 x20 x21 x22 x23 x24 x25 x26 x27 x28 x29 x30 x31 x32 x33 x34 x35, fe

Fixed-effects (within) regression              Number of obs   =       2592
Group variable: panel                        Number of groups  =         24

R-sq:  within = 0.3434                      Obs per group: min =        108
        between = 0.4610                      avg           =       108.0
        overall = 0.3683                      max           =        108

corr(u_i, Xb) = 0.4417                      F(35,2533)       =       37.84
                                                Prob > F         =       0.0000
```

logproporc~n	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
x1	.0008573	.0007278	1.18	0.239	-.0005698	.0022844
x2	1.268056	.1150882	11.02	0.000	1.04238	1.493733
x3	.181577	.0893815	2.03	0.042	.0063087	.3568453
x4	-.6156916	.0816541	-7.54	0.000	-.7758073	-.4555759
x5	-.000784	.0005648	-1.39	0.165	-.0018914	.0003234
x6	-.0013443	.0007945	-1.69	0.091	-.0029023	.0002137
x7	-.0014488	.0006299	-2.30	0.022	-.002684	-.0002137
x8	.4491445	.1321314	3.40	0.001	.190048	.7082411
x9	.0089578	.0015729	5.70	0.000	.0058736	.012042
x10	.0106603	.0012074	8.83	0.000	.0082927	.0130279
x11	.0071512	.0009243	7.74	0.000	.0053387	.0089638
x12	-.0014398	.0005892	-2.44	0.015	-.0025952	-.0002844
x13	.0048143	.0006503	7.40	0.000	.0035391	.0060894
x14	.0034161	.0006005	5.69	0.000	.0022385	.0045936
x15	-.4152261	.1229778	-3.38	0.001	-.6563734	-.1740788
x16	-.0000157	.0009334	-0.02	0.987	-.001846	.0018145
x17	.0014775	.0008628	1.71	0.087	-.0002144	.0031694
x18	.0010332	.0006516	1.59	0.113	-.0002445	.0023108
x19	-.0014244	.0012671	-1.12	0.261	-.0039091	.0010602
x20	.0024166	.0005323	4.54	0.000	.0013729	.0034604
x21	.0006979	.0004896	1.43	0.154	-.0002621	.0016579
x22	.0536691	.0469533	1.14	0.253	-.0384016	.1457399
x23	.0048576	.0024728	1.96	0.050	8.74e-06	.0097064
x24	.0001171	.0007267	0.16	0.872	-.0013079	.0015421
x25	.0012053	.0011286	1.07	0.286	-.0010079	.0034184
x26	.0003748	.0005558	0.67	0.500	-.000715	.0014646
x27	-.0013127	.0005726	-2.29	0.022	-.0024354	-.0001899
x28	-.0036517	.0036646	-1.00	0.319	-.0108376	.0035342
x29	-.0365664	.0918723	-0.40	0.691	-.216719	.1435861
x30	-.0028863	.0016392	-1.76	0.078	-.0061007	.0003281
x31	.0014887	.0007546	1.97	0.049	9.03e-06	.0029683
x32	-.0010325	.0004986	-2.07	0.038	-.0020102	-.0000549
x33	-.0007597	.0056027	-0.14	0.892	-.011746	.0102266
x34	-.0006315	.0002952	-2.14	0.033	-.0012104	-.0000525
x35	-.1252656	.0494131	-2.54	0.011	-.2221598	-.0283713
_cons	-3.565787	.2380975	-14.98	0.000	-4.032673	-3.098901
sigma_u	1.0471448					
sigma_e	.23171983					
rho	.95331787	(fraction of variance due to u_i)				
F test that all u_i=0: F(23, 2533) = 488.35 Prob > F = 0.0000						

Hay que hacer notar las grandes discrepancias que existen en ambos modelos en los valores del coeficiente de determinación, ya que en el modelo de efectos agrupados alcanza el 79,95 % y en el de efectos fijos es del 36,83 %. A primera vista, y si nos fijáramos en este coeficiente como criterio de selección de modelos parece que el primer modelo tiene una mayor capacidad explicativa que el segundo.

Por esta razón vamos a utilizar el contraste propuesto por Hsiao (2002,15).

En este contraste la hipótesis a contrastar es:  $H_4: \alpha_1 = \alpha_2 = \dots \alpha_N$  esto es, dado  $\beta_1 = \beta_2 = \dots \beta_N$

asumiendo que las  $\beta$  son las misma para los dos modelos, se trata de ver si los puntos de corte son los mismo o existe un punto de corte distinto para cada individuo que compone el panel. Esto es, la estimación adecuada es un modelo de efectos fijos (con puntos de corte distintos por individuo) frente a un modelo de efectos conjuntos (un solo punto de corte para todos los individuos).

El estadístico de contraste es:  $F_4 = \frac{\frac{S_3 - S_2}{(N-1)}}{\frac{S_2}{N(T-1) - K}}$  rechazando la hipótesis nula si

$$F_4 \geq F_{(N-1), N(T-1)-K}^{(\alpha)}$$

Para este caso concreto los valores del estadístico son

- $S_3 = 739,097704$  que aparece en la salida por defectos de la regresión de MCO agrupados o que también puede calcularse guardando los residuos del modelo, elevándolos al cuadrado y sumándolos.
- $S_2$  este valor no aparece por defecto en la regresión del modelo  $y_{it} = \alpha_i + \beta' x_{it} + u_{it}$  por efectos fijos. En este caso el valor concreto es 136,0071003
- $N$  número de individuos que componen el panel y en este caso es 24
- $T$  número de periodos temporales considerados; 108
- $K$  número de variables explicativas consideradas; en este caso y con la constante 36



Sustituyendo en la expresión del estadístico

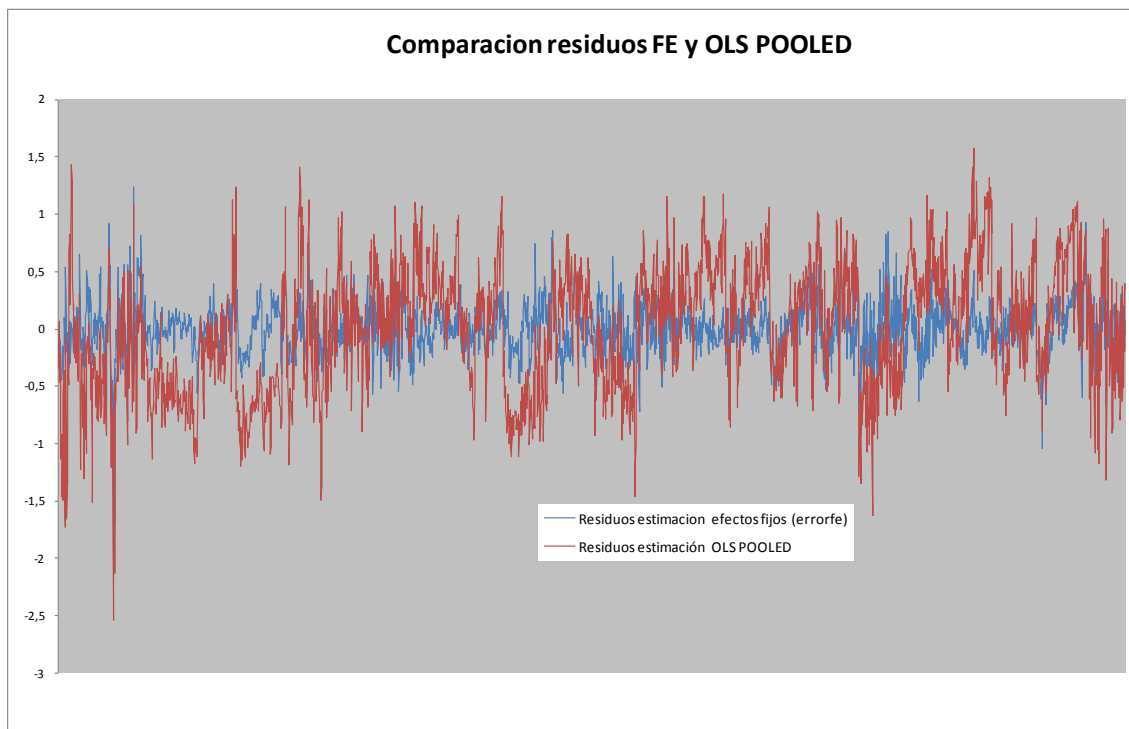
$$F_4 = \frac{\frac{739,097704 - 136,0071003}{(24-1)}}{\frac{136,0071003}{24(108-1) - 36}} = 488,1539916$$

Comparándolo con el valor en tablas tenemos:

$488,1539916 = F_4 \geq F_{(N-1), N(T-1)-K}^{(\alpha)} = F_{(24-1), 24(108-1)-36}^{(5\%)} = F_{23, 2532}^{(5\%)} = 0.971425$  por lo tanto rechazamos la hipótesis nula. En términos de *pvalor*, el *pvalor* asociado al estadístico es 0, inferior a cualquier nivel de significación que pudiéramos elegir, por lo tanto, la hipótesis nula se rechazaría.

En definitiva, podemos afirmar que el modelo  $y_{it} = \alpha + \mathbf{x}'_{it}\beta + u_{it}$  no es el adecuado comparándolo con el modelo  $y_{it} = \alpha_i + \beta'x_{it} + u_{it}$ . Ya que el estimador utilizado resulta ser INCONSISTENTE.

Una forma alternativa de verlos es graficando conjuntamente los residuos producidos por ambos modelos.



Puede observarse que los residuos de la estimación MCO Agrupados presentan una mayor variabilidad que los de la estimación de efectos fijos, por lo tanto sus cuadrados deben ser mayores que los cuadrados de los residuos de la estimación de efectos fijos, cosa que se demuestra en el contraste anteriormente propuesto. Este gráfico no es más que una forma alternativa de comprobar la eficacia del contraste.

Desde el punto de vista de marketing, la aplicación del contraste de Hsiao (2002), además de contrastar una de las hipótesis de partida contenidas en esta tesis, supone admitir que NO hay un único modelo de gestión para todos los formatos y establecimientos; sino que es necesario realizar adaptaciones a las particularidades de cada establecimiento y formato.

En este momento, desconocemos aún cuál será la profundidad de esta heterogeneidad; esto es, si la heterogeneidad presente en cada individuo podemos recogerla en una constante específica para él o por el contrario esta heterogeneidad será de naturaleza más amplia y será preciso recogerla en un modelo para cada

individuo. O quizás quedarnos en un punto intermedio; esto es, un modelo para cada formato o un modelo para cada tipo de establecimiento.

Cuestiones que trataremos de responder en los epígrafes siguientes.

### 5.4.3- MODELO DE EFECTOS INDIVIDUALES ESPECÍFICOS

En este caso el modelo  $y_{it} = \alpha_i + \mathbf{x}_{it}'\beta + u_{it}$  permite que cada individuo (unidad de sección cruzada) tenga un intercepto diferente, aunque la pendiente siga siendo la misma para todos los individuos; en este modelo asumimos que  $u_{it}$  está iid en  $i$  y  $t$ .

Los  $\alpha_i$  son variables aleatorias que recogen la heterogeneidad no observada

Una variante de este modelo es tratar a  $\alpha_i$  como una variable aleatoria inobservada que potencialmente está correlacionada con los  $\mathbf{x}_{it}$  regresores observados. Esta variante recibe el nombre de modelo de efectos fijos (EF). Si los efectos fijos están presentes y los  $\alpha_i$  están correlacionados con las  $\mathbf{x}_{it}$  un buen número de estimadores resultan ser inconsistentes.

Otra variante del modelo  $y_{it} = \alpha_i + \mathbf{x}_{it}'\beta + u_{it}$  es asumir que los  $\alpha_i$ , que representan los efectos individuales inobservable son variables aleatorias que se distribuyen independientemente de las regresores. Este modelo es el modelo de efectos aleatorios (EA), que normalmente establece los supuestos adicionales:

$$\begin{aligned}\alpha_i &\sim [\alpha, \sigma_\alpha^2] \\ u_{it} &\sim [0, \sigma_u^2]\end{aligned}$$

La forma de proceder será estimar ambos modelos y después elegir el más adecuado para nuestro análisis.

#### 5.4.3.1.- Estimador Within o de Efectos Fijos (EF)

Este estimador se obtiene restando al modelo original el modelo de promedio temporal

$$\bar{y}_i = \alpha_i + \bar{\mathbf{x}}_i' \beta + \bar{u}_i$$

Entonces el modelo transformado  $y_{it} - \bar{y}_i = (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)' \beta + (u_{it} - \bar{u}_i)$  elimina el efecto individual  $\alpha_i$  junto con todos los regresores invariantes en el tiempo, ya que  $\mathbf{x}_{it} - \bar{\mathbf{x}}_i = 0$  si  $\mathbf{x}_{it} = \mathbf{x}_i$  para todo  $t$ .

El estimador de Efectos Fijos resulta de aplicar Mínimos Cuadrados Ordinarios sobre el modelo transformado  $y_{it} - \bar{y}_i = (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)' \beta + (u_{it} - \bar{u}_i)$

Los efectos individuales  $\alpha_i$  pueden estimarse mediante

$$\hat{\alpha}_i = \bar{y}_i - \bar{\mathbf{x}}_i' \hat{\beta}_{EF}, \quad i = 1, \dots, N$$

El estimador  $\hat{\alpha}_i$  es insesgado y consistente si  $T \rightarrow \infty$

El estimador de  $\beta$  es consistente si  $p \lim \frac{1}{NT} \sum_i \sum_t (\mathbf{x}_{it} - \bar{\mathbf{x}}_i)(u_{it} - \bar{u}_i) = 0$ , Esto debería ocurrir tanto si  $N \rightarrow \infty$  o  $T \rightarrow \infty$  y  $E[u_{it} - \bar{u}_i / \mathbf{x}_{it} - \bar{\mathbf{x}}_i] = 0$ . Esta condición supone excluir de la estimación de Efectos Fijos la posibilidad de incluir como regresores variables exógenas retardadas.

En nuestro caso la estimación de Efectos Fijos produce los resultados siguientes:

```
. xtreg logproporción x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 x13 x14 x15 x16 x17 x18 x19 x20 x21 x22 x23 x24 x25 x26 x27 x28 x29 x30 x31 x32 x33 x34 x35, fe
```

Fixed-effects (within) regression  
Group variable: panel

Number of obs = 2592  
Number of groups = 24

R-sq: within = 0.3434  
between = 0.4610  
overall = 0.3683

Obs per group: min = 108  
avg = 108.0  
max = 108

corr(u\_i, xb) = 0.4417

F(35,2533) = 37.84  
Prob > F = 0.0000

logproporc~n	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
x1	.0008573	.0007278	1.18	0.239	-.0005698	.0022844
x2	1.268056	.1150882	11.02	0.000	1.04238	1.493733
x3	.181577	.0893815	2.03	0.042	.0063087	.3568453
x4	-.6156916	.0816541	-7.54	0.000	-.7758073	-.4555759
x5	-.000784	.0005648	-1.39	0.165	-.0018914	.0003234
x6	-.0013443	.0007945	-1.69	0.091	-.0029023	.0002137
x7	-.0014488	.0006299	-2.30	0.022	-.002684	-.0002137
x8	.4491445	.1321314	3.40	0.001	.190048	.7082411
x9	.0089578	.0015729	5.70	0.000	.0058736	.012042
x10	.0106603	.0012074	8.83	0.000	.0082927	.0130279
x11	.0071512	.0009243	7.74	0.000	.0053387	.0089638
x12	-.0014398	.0005892	-2.44	0.015	-.0025952	-.0002844
x13	.0048143	.0006503	7.40	0.000	.0035391	.0060894
x14	.0034161	.0006005	5.69	0.000	.0022385	.0045936
x15	-.4152261	.1229778	-3.38	0.001	-.6563734	-.1740788
x16	-.0000157	.0009334	-0.02	0.987	-.001846	.0018145
x17	.0014775	.0008628	1.71	0.087	-.0002144	.0031694
x18	.0010332	.0006516	1.59	0.113	-.0002445	.0023108
x19	-.0014244	.0012671	-1.12	0.261	-.0039091	.0010602
x20	.0024166	.0005323	4.54	0.000	.0013729	.0034604
x21	.0006979	.0004896	1.43	0.154	-.0002621	.0016579
x22	.0536691	.0469533	1.14	0.253	-.0384016	.1457399
x23	.0048576	.0024728	1.96	0.050	8.74e-06	.0097064
x24	.0001171	.0007267	0.16	0.872	-.0013079	.0015421
x25	.0012053	.0011286	1.07	0.286	-.0010079	.0034184
x26	.0003748	.0005558	0.67	0.500	-.000715	.0014646
x27	-.0013127	.0005726	-2.29	0.022	-.0024354	-.0001899
x28	-.0036517	.0036646	-1.00	0.319	-.0108376	.0035342
x29	-.0365664	.0918723	-0.40	0.691	-.216719	.1435861
x30	-.0028863	.0016392	-1.76	0.078	-.0061007	.0003281
x31	.0014887	.0007546	1.97	0.049	9.03e-06	.0029683
x32	-.0010325	.0004986	-2.07	0.038	-.0020102	-.0000549
x33	-.0007597	.0056027	-0.14	0.892	-.011746	.0102266
x34	-.0006315	.0002952	-2.14	0.033	-.0012104	-.0000525
x35	-.1252656	.0494131	-2.54	0.011	-.2221598	-.0283713
_cons	-3.565787	.2380975	-14.98	0.000	-4.032673	-3.098901
sigma_u	1.0471448					
sigma_e	.23171983					
rho	.95331787	(fraction of variance due to u_i)				

F test that all u\_i=0: F(23, 2533) = 488.35 Prob > F = 0.0000

En estos resultados puede observarse el contraste F para comprobar si las constantes individuales son 0 o no. A la luz de los datos puede observarse que no aceptamos la hipótesis nula de que todas las constantes individuales son 0.

### 5.4.3.2.- Estimador de Efectos Aleatorios (EA)

Este estimador parte del mismo modelo  $y_{it} = \alpha_i + \mathbf{x}'_{it}\beta + u_{it}$  pero asumiendo que los interceptos  $\alpha_i$  son variables aleatorias que se distribuyen independientemente de los regresores; además vamos a suponer que  $\alpha_i$  y  $u_{it}$  son independientes y están idénticamente distribuidos.

$$\alpha_i \sim [\alpha, \sigma_\alpha^2]$$

$$u_{it} \sim [0, \sigma_u^2]$$

El estimador de Efectos Aleatorios puede calcularse como el estimador MCO sobre el modelo transformado  $y_{it} - \hat{\lambda}\bar{y}_i = (1 - \hat{\lambda})\mu + (x_{it} - \hat{\lambda}\bar{x}_i)' \beta + v_{it}$  donde  $v_{it} = (1 - \hat{\lambda})\alpha_i + (u_{it} - \hat{\lambda}\bar{u}_i)$  es asintóticamente independiente y está idénticamente distribuido. Además  $\hat{\lambda}$  es el estimador consistente ya que  $\lambda = 1 - \frac{\sigma_u}{\sqrt{\sigma_u^2 + T\sigma_\alpha^2}}$

```
. xtreg logproporcion x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 x13 x14 x15 x16 x17 x
> 18 x19 x20 x21 x22 x23 x24 x25 x26 x27 x28 x29 x30 x31 x32 x33 x34 x35, re theta
```

Random-effects GLS regression                      Number of obs        =        2592  
Group variable: panel                      Number of groups     =        24

R-sq:    within = 0.1385                      obs per group: min =        108  
          between = 0.9072                                     avg =        108.0  
          overall = 0.7995                                     max =        108

Random effects u\_i ~ Gaussian                      Wald chi2(35)        =    10190.23  
corr(u\_i, X)        = 0 (assumed)                      Prob > chi2        =        0.0000  
theta                = 0

logproporc~n	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
x1	-.0025274	.0005887	-4.29	0.000	-.0036812	-.0013736
x2	-.2432966	.1930804	-1.26	0.208	-.6217271	.1351339
x3	1.636362	.0909354	17.99	0.000	1.458132	1.814593
x4	.9255853	.1405039	6.59	0.000	.6502027	1.200968
x5	-.0066885	.0006594	-10.14	0.000	-.0079809	-.0053961
x6	.0010027	.0005487	1.83	0.068	-.0000727	.0020782
x7	.0054928	.0005899	9.31	0.000	.0043366	.006649
x8	.3817528	.2876013	1.33	0.184	-.1819354	.945441
x9	.0379131	.003444	11.01	0.000	.0311631	.0446631
x10	.0266081	.0026018	10.23	0.000	.0215086	.0317076
x11	.0150357	.0020847	7.21	0.000	.0109498	.0191216
x12	.0145386	.0011445	12.70	0.000	.0122954	.0167818
x13	.008858	.0013905	6.37	0.000	.0061326	.0115833
x14	.0146833	.0011131	13.19	0.000	.0125016	.016865
x15	.1188997	.2688868	0.44	0.658	-.4081087	.6459081
x16	.0084599	.0021228	3.99	0.000	.0042994	.0126204
x17	.0061703	.0019336	3.19	0.001	.0023804	.0099602
x18	-.0056199	.0014738	-3.81	0.000	-.0085085	-.0027314
x19	-.0439024	.0017865	-24.57	0.000	-.0474039	-.0404009
x20	-.0008473	.0012023	-0.70	0.481	-.0032038	.0015092
x21	-.0012597	.001082	-1.16	0.244	-.0033804	.0008609
x22	.7626727	.1002483	7.61	0.000	.5661896	.9591557
x23	-.0140135	.005317	-2.64	0.008	-.0244346	-.0035924
x24	.004096	.0016301	2.51	0.012	.0009011	.0072909
x25	.0065635	.0025745	2.55	0.011	.0015175	.0116095
x26	.0036847	.0009458	3.90	0.000	.001831	.0055383
x27	-.0011681	.0013005	-0.90	0.369	-.0037171	.0013808
x28	.0073334	.0082792	0.89	0.376	-.0088935	.0235604
x29	.1946226	.2082374	0.93	0.350	-.2135153	.6027604
x30	.0061371	.0037738	1.63	0.104	-.0012595	.0135337
x31	.0094547	.0016825	5.62	0.000	.006157	.0127524
x32	-.0008817	.001117	-0.79	0.430	-.0030709	.0013075
x33	-.0135871	.0126311	-1.08	0.282	-.0383437	.0111694
x34	.0008247	.0006742	1.22	0.221	-.0004966	.0021461
x35	-.1873359	.1136026	-1.65	0.099	-.4099929	.0353212
_cons	-5.390621	.4744018	-11.36	0.000	-6.320431	-4.460811
sigma_u	0					
sigma_e	.23171983					
rho	0	(fraction of variance due to u_i)				

Las diferencias entre ambas estimaciones son bastante notables, pero aún es demasiado pronto para decidir con cuál de ellas debemos quedarnos. Este paso lo vamos a dar con la ayuda del Test de Hausman (1978), que está diseñado para situaciones como esta.



. hausman fe re

	Coefficients			
	(b) fe	(b) re	(b-B) Difference	sqrt(diag(V_b-V_B)) S.E.
x1	.0008573	-.0025274	.0033847	.0004279
x2	1.268056	-.2432966	1.511353	.
x3	.181577	1.636362	-1.454785	.
x4	-.6156916	.9255853	-1.541277	.
x5	-.000784	-.0066885	.0059045	.
x6	-.0013443	.0010027	-.002347	.0005746
x7	-.0014488	.0054928	-.0069417	.0002209
x8	.4491445	.3817528	.0673917	.
x9	.0089578	.0379131	-.0289553	.
x10	.0106603	.0266081	-.0159478	.
x11	.0071512	.0150357	-.0078845	.
x12	-.0014398	.0145386	-.0159784	.
x13	.0048143	.008858	-.0040437	.
x14	.0034161	.0146833	-.0112672	.
x15	-.4152261	.1188997	-.5341258	.
x16	-.0000157	.0084599	-.0084756	.
x17	.0014775	.0061703	-.0046928	.
x18	.0010332	-.0056199	.0066531	.
x19	-.0014244	-.0439024	.042478	.
x20	.0024166	-.0008473	.003264	.
x21	.0006979	-.0012597	.0019576	.
x22	.0536691	.7626727	-.7090036	.
x23	.0048576	-.0140135	.0188711	.
x24	.0001171	.004096	-.0039788	.
x25	.0012053	.0065635	-.0053583	.
x26	.0003748	.0036847	-.0033099	.
x27	-.0013127	-.0011681	-.0001445	.
x28	-.0036517	.0073334	-.0109852	.
x29	-.0365664	.1946226	-.231189	.
x30	-.0028863	.0061371	-.0090234	.
x31	.0014887	.0094547	-.007966	.
x32	-.0010325	-.0008817	-.0001508	.
x33	-.0007597	-.0135871	.0128274	.
x34	-.0006315	.0008247	-.0014562	.
x35	-.1252656	-.1873359	.0620703	.

b = consistent under Ho and Ha; obtained from xtreg  
B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(35) = (b-B)'[(V\_b-V\_B)^(-1)](b-B)  
= 1255.85  
Prob>chi2 = 0.0000  
(V\_b-V\_B is not positive definite)

En este contraste, como establecíamos en el capítulo 3, comparamos la diferencia entre las estimaciones de Efectos Fijos (EF) y Efectos Aleatorios (EA) con el objetivo de comprobar si esas diferencias son sistemáticas o no. De tal forma que si aceptamos la hipótesis nula (diferencias no sistemáticas) el estimador de Efectos Aleatorios sería el preferido, ya que sería eficiente, cosa que no sucede con el estimador de Efectos Fijos.

A la vista del resultado rechazamos la hipótesis nula pues el p valor es inferior al nivel de significación; por tanto el estimador que debemos utilizar es el estimador de Efectos Fijos: Ahora bien, este contraste no es definitivo ya que el Test de Hausman estándar que se ha utilizado requiere que el estimador de Efectos Aleatorios (EA) sea eficiente. Esto significa que  $\alpha_i$  y  $u_{it}$  estén independiente e idénticamente distribuidos (*i.i.d.*) pero esta hipótesis de partida resulta invalidada si se produce una

diferencia sustancial entre los errores estándar y los errores estándar robustos por grupos. Por tanto parece necesaria una versión robusta del test de Hausman.

Lo más sencillo es testar  $H_0 : \gamma = 0$  en la regresión auxiliar de MCO

$$(y_{it} - \hat{\theta}\bar{y}_i) = (1 - \hat{\theta})\alpha + (x_{it} - \hat{\theta}\bar{x}_i)' \beta + (x_{1it} - \hat{\theta}\bar{x}_{1i})' \gamma + u_{it}$$

Donde  $x_{1i}$  hace referencia exclusivamente a los regresores que varían en el tiempo.

Un test de Wald de  $\gamma = 0$  es asintóticamente equivalente al test de Hausman estándar cuando el estimador de efectos aleatorios (EA) es completamente eficiente bajo  $H_0$ .

En el caso que el estimador EA no sea eficiente, Wooldridge (2002, 262) propone un test de Wald usando errores estándar robustos por grupos. Esto es lo que se ha hecho.

En el cuadro que tenemos a continuación viene reflejado el resultado del método propuesto por Wooldridge, así como las instrucciones para realizarlo en Stata.

```

. regress cdllogprop cdx1 cdx2 cdx3 cdx4 cdx5 cdx6 cdx7 cdx8 cdx9 cdx10 cdx11 c
v dx12 cdx13 cdx14 cdx15 cdx16 cdx17 cdx18 cdx19 cdx20 cdx21 cdx22 cdx23 cdx24
v cdx25 cdx26 cdx27 cdx28 cdx29 cdx30 cdx31 cdx32 cdx33 cdx34 cdx35 dx1 dx2 dx3
v dx4 dx5 dx6 dx7 dx8 dx9 dx10 dx11 dx12 dx13 dx14 dx15 dx16 dx17 dx18 dx19 dx
v 20 dx21 dx22 dx23 dx24 dx25 dx26 dx27 dx28 dx29 dx30 dx31 dx32 dx33 dx34 dx35
v , vce(cluster panel)
note: cdx3 omitted because of collinearity
note: cdx4 omitted because of collinearity
note: cdx15 omitted because of collinearity
note: cdx22 omitted because of collinearity
note: dx2 omitted because of collinearity
note: dx8 omitted because of collinearity
note: dx28 omitted because of collinearity
note: dx29 omitted because of collinearity
note: dx30 omitted because of collinearity
note: dx33 omitted because of collinearity
note: dx35 omitted because of collinearity

```

Linear regression

```

Number of obs = 2592
F( 23, 23) = .
Prob > F = .
R-squared = 0.9874
Root MSE = .13526

```

(Std. Err. adjusted for 24 clusters in panel)

cdllogprop	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
cdx1	-.0024069	.0009866	-2.44	0.023	-.0044479	-.000366
cdx2	.1422606	.1874998	0.76	0.456	-.2456122	.5301334
cdx3	(omitted)					
cdx4	(omitted)					
cdx5	.0115492	.0012523	9.22	0.000	.0089586	.0141399
cdx6	-.0334394	.0025802	-12.96	0.000	-.0387769	-.0281018
cdx7	.0388934	.0010926	35.60	0.000	.0366333	.0411536
cdx8	.0868004	.1055727	0.82	0.419	-.1315934	.3051942
cdx9	-.0171221	.0817097	-0.21	0.836	-.1861515	.1519073
cdx10	-.5052261	.0391264	-12.91	0.000	-.5861653	-.424287
cdx11	.4891985	.0504018	9.71	0.000	.3849345	.5934626
cdx12	-.0576635	.0028382	-20.32	0.000	-.0635348	-.0517922
cdx13	.0794766	.0258273	3.08	0.005	.0260487	.1329045
cdx14	.1039072	.0065834	15.78	0.000	.0902884	.117526
cdx15	(omitted)					
cdx16	.7393778	.0161629	45.75	0.000	.7059422	.7728134
cdx17	-.0506464	.0065551	-7.73	0.000	-.0641981	-.0370947
cdx18	-.5467414	.0096991	-56.37	0.000	-.5668055	-.5266773
cdx19	.0537749	.0076301	7.05	0.000	.0379908	.0695591
cdx20	.2188421	.0111136	19.69	0.000	.1958519	.2418322
cdx21	.0930164	.0168922	5.51	0.000	.0580722	.1279605
cdx22	(omitted)					
cdx23	.6331288	.015919	39.77	0.000	.6001979	.6660597
cdx24	-.1498807	.0203017	-7.38	0.000	-.1918778	-.1078835
cdx25	.1264121	.0069963	18.07	0.000	.1119392	.1408851
cdx26	.0033447	.0047281	0.71	0.486	-.0064362	.0131256
cdx27	.1002041	.0079219	12.65	0.000	.0838164	.1165918
cdx28	.0065312	.0025211	2.59	0.016	.0013158	.0117466
cdx29	-.0359031	.0904932	-0.40	0.695	-.2231026	.1512965
cdx30	-.0000563	.0005537	-0.10	0.920	-.0012017	.0010891
cdx31	.253106	.0095242	26.57	0.000	.2334037	.2728084
cdx32	-.1331132	.0130483	-10.20	0.000	-.1601057	-.1061207
cdx33	.0013396	.0007574	1.77	0.090	-.0002272	.0029063
cdx34	.0845512	.0067058	12.61	0.000	.0706792	.0984233
cdx35	.027509	.0319844	0.86	0.399	-.0386558	.0936739
dx1	8.151611	.7149282	11.40	0.000	6.672669	9.630552
dx2	(omitted)					
dx3	-.1383116	.0571372	-2.42	0.024	-.256509	-.0201143
dx4	-.1138579	.0530734	-2.15	0.043	-.2236487	-.0040672
dx5	-.0125164	.0014769	-8.47	0.000	-.0155716	-.0094612
dx6	.0337755	.0028293	11.94	0.000	.0279227	.0396283
dx7	-.0377319	.0016896	-22.33	0.000	-.041227	-.0342367
dx8	(omitted)					
dx9	.0149377	.0823027	0.18	0.858	-.1553183	.1851938
dx10	.5027172	.0390397	12.88	0.000	.4219574	.583477
dx11	-.4909138	.0503032	-9.76	0.000	-.5949739	-.3868537
dx12	.0575586	.0027074	21.26	0.000	.0519579	.0631593
dx13	-.0792987	.0261511	-3.03	0.006	-.1333964	-.025201
dx14	-.1044761	.00686	-15.23	0.000	-.1186672	-.0902851
dx15	-.1433158	.089163	-1.61	0.122	-.3277635	.041132
dx16	-.7389045	.0159626	-46.29	0.000	-.7719257	-.7058833
dx17	.0511914	.006755	7.58	0.000	.0372177	.0651651
dx18	.5466199	.0096224	56.81	0.000	.5267144	.5665254
dx19	-.0550903	.0078295	-7.04	0.000	-.0712868	-.0388937
dx20	-.2183646	.011188	-19.52	0.000	-.2415087	-.1952205
dx21	-.093402	.0167758	-5.57	0.000	-.1281053	-.0586987
dx22	.0654427	.0473013	1.38	0.180	-.0324074	.1632928
dx23	-.6344776	.0159583	-39.76	0.000	-.66749	-.6014653
dx24	.149549	.0199986	7.48	0.000	.1081787	.1909193
dx25	-.1259655	.0069838	-18.04	0.000	-.1404126	-.1115183
dx26	-.0029304	.0047148	-0.62	0.540	-.0126837	.0068228
dx27	-.1006108	.0079354	-12.68	0.000	-.1170265	-.084195
dx28	(omitted)					
dx29	(omitted)					
dx30	(omitted)					
dx31	-.2537052	.009385	-27.03	0.000	-.2731195	-.2342909
dx32	.1328429	.0129702	10.24	0.000	.1060119	.1596739
dx33	(omitted)					
dx34	-.0846331	.0066919	-12.65	0.000	-.0984764	-.0707899
dx35	(omitted)					
_cons	-7.602239	.2135943	-35.59	0.000	-8.044093	-7.160386

$Cdx$  hace referencia a las variables que han sufrido la transformación  $(x_{it} - \hat{\theta}\bar{x}_i)$  y  $dx$  se refiere a las variables que varían en el tiempo y experimentan la transformación  $(x_{lit} - \hat{\theta}\bar{x}_{li})$ . Por otra parte, la estimación por MCO se ha realizado activando la opción de *vce (cluster panel)* que permite la estimación robusta por grupos.

Los resultados del test de Wald realizado a continuación son los siguientes:

```
. test  dx1 dx2 dx3 dx4 dx5 dx6 dx7 dx8 dx9 dx10 dx11 dx12 dx13 dx14 dx15 dx16
> dx17 dx18 dx19 dx20 dx21 dx22 dx23 dx24 dx25 dx26 dx27 dx28 dx29 dx30 dx31 dx
> 32 dx33 dx34 dx35

( 1)  dx1 = 0
( 2)  o.dx2 = 0
( 3)  dx3 = 0
( 4)  dx4 = 0
( 5)  dx5 = 0
( 6)  dx6 = 0
( 7)  dx7 = 0
( 8)  o.dx8 = 0
( 9)  dx9 = 0
(10)  dx10 = 0
(11)  dx11 = 0
(12)  dx12 = 0
(13)  dx13 = 0
(14)  dx14 = 0
(15)  dx15 = 0
(16)  dx16 = 0
(17)  dx17 = 0
(18)  dx18 = 0
(19)  dx19 = 0
(20)  dx20 = 0
(21)  dx21 = 0
(22)  dx22 = 0
(23)  dx23 = 0
(24)  dx24 = 0
(25)  dx25 = 0
(26)  dx26 = 0
(27)  dx27 = 0
(28)  o.dx28 = 0
(29)  o.dx29 = 0
(30)  o.dx30 = 0
(31)  dx31 = 0
(32)  dx32 = 0
(33)  o.dx33 = 0
(34)  dx34 = 0
(35)  o.dx35 = 0
      Constraint 2 dropped
      Constraint 8 dropped
      Constraint 14 dropped
      Constraint 17 dropped
      Constraint 20 dropped
      Constraint 27 dropped
      Constraint 28 dropped
      Constraint 29 dropped
      Constraint 30 dropped
      Constraint 33 dropped
      Constraint 34 dropped
      Constraint 35 dropped

      F( 23,      23) = 1.5e+06
          Prob > F =      0.0000
```

A la vista de los resultados se rechaza la hipótesis de  $H_0:\gamma=0$  por lo que concluimos que el modelo de Efectos Aleatorios no es el apropiado.

Una cosa más a añadir respecto al modelo de Efectos Aleatorios. Como ya hemos establecido anteriormente, este modelo supone realizar la transformación

$y_{it} - \hat{\lambda} \bar{y}_i = (1 - \hat{\lambda})\mu + (x_{it} - \hat{\lambda} \bar{x}_i)' \beta + v_{it}$  respecto al modelo original. Donde  $\hat{\lambda}$  es el estimador consistente ya que  $\lambda = 1 - \frac{\sigma_u}{\sqrt{\sigma_u^2 + T\sigma_\alpha^2}}$  y el que los estimadores de cada

uno de los parámetros son:

$$\hat{\sigma}_u^2 = \frac{1}{N(T-1) - K} \sum_i \sum_t \left[ (y_{it} - \bar{y}_i) - (x_{it} - \bar{x}_i)' \beta_{FE} \right]^2$$

$$\hat{\sigma}_\alpha^2 = \frac{1}{N - (K+1)} \sum_i \left( \bar{y}_i - \hat{\mu}_{Between} - \bar{x}_i' \hat{\beta}_{Between} \right)^2 - \frac{1}{T} \hat{\sigma}_u^2$$

Este último estimador de la varianza puede ser negativo, en casos como este los programas hacen que  $\hat{\sigma}_\alpha^2 = 0$ , y por tanto  $\hat{\lambda} = 0$ , por lo que la estimación de Efectos Aleatorios coincide con la MCO Agrupados, esto es lo que sucede en este caso, los parámetros estimados por Efectos Aleatorios y por MCO Agrupados son iguales.

En este momento parece claro que el modelo más adecuado para nuestro objetivo es el de Efectos Fijos; pero no estaría de más responder a la cuestión de si es preferible un modelo global para todos los establecimientos y formatos o un modelo para cada establecimiento y formato.

La respuesta a esta pregunta está relacionada con una importante cuestión en marketing. Existe una sola política comercial o bien podemos hablar de una política comercial diferente en función del formato y del tipo de establecimiento donde se comercializa el producto.

Para responderla vamos a utilizar una versión distinta del test propuesto por Hsiao (2002, 15 - 21). En este caso la hipótesis a plantear es:  $H_1 : \beta_1 = \beta_2 = \dots \beta_N$  Aceptar

$H_1$  significa que el modelo adecuado es  $y_{it} = \alpha_i + \beta' x_{it} + u_{it}$  mientras que rechazarla supone admitir que este modelo no es el adecuado y por tanto, tenemos un modelo para cada individuo.

El estadístico de contraste que se utiliza es:  $F_1 = \frac{\frac{S_2 - S_1}{(N-1)K}}{\frac{S_1}{NT - N(K+1)}}$ ; rechazando la

hipótesis si  $F_1 \geq F_{(N-1)K, NT-N(K+1)}^{(\alpha)}$

Para este caso concreto los valores del estadístico son

- $S_2 = 136,0071003$  es la suma de residuos al cuadrado de la regresión por Efectos Fijos.
- $S_1$  es la suma de los residuos al cuadrado. Para calcularlo hacemos una regresión para cada individuo y sumamos los residuos al cuadrado de todas y cada una de las regresiones realizadas (una por individuo). Su valor en este caso es: 64,90194318
- $N$  número de individuos que componen el panel y en este caso es 24
- $T$  número de periodos temporales considerados; 108
- $K$  número de variables explicativas consideradas; en este caso y con la constante 36

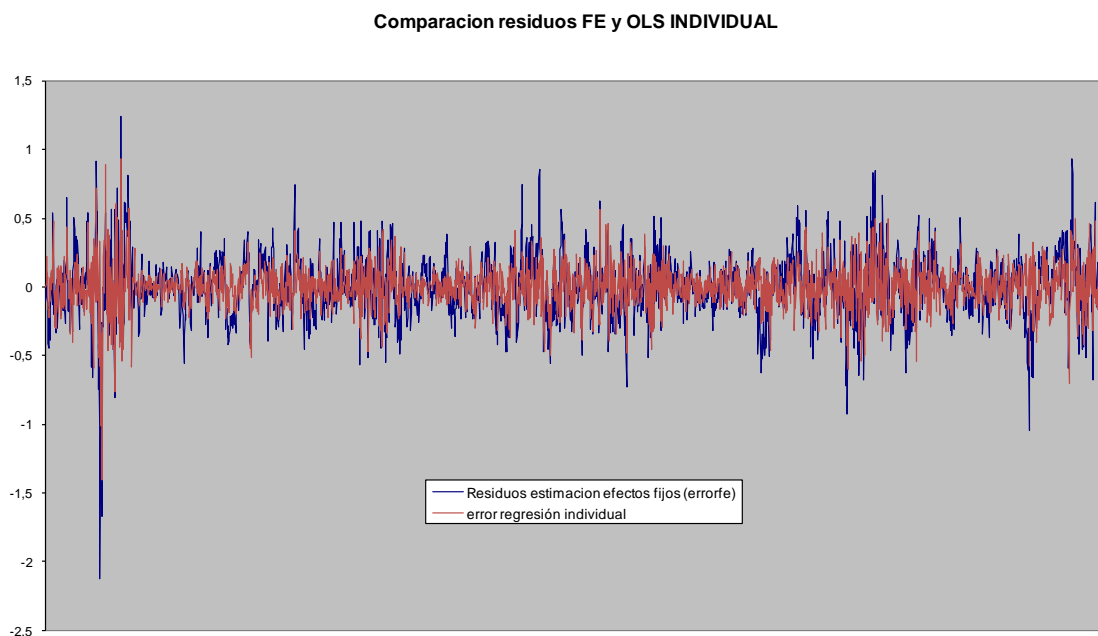
Sustituyendo en la expresión del estadístico

$$F_1 = \frac{\frac{136,0071003 - 64,90194318}{(24-1) \cdot 36}}{\frac{64,90194318}{24 \cdot 108 - 24 \cdot (36+1)}} = 903,734898$$

Comparándolo con el valor en tablas tenemos:

$903,734898 = F_1 \geq F_{(N-1)K, NT-N(K+1)}^{(\alpha)} = F_{(24-1) \cdot 36, 24 \cdot 108 - 24 \cdot (36+1)}^{(5\%)} = 1.10254$  por lo tanto rechazamos la hipótesis nula. Esto es, no existe un único modelo común a todos los

individuos; que en términos más de marketing supone admitir que no existe una única política comercial para todos los clientes. Todavía resulta un poco prematuro saber qué tipo de política comercial existe o dicho de otra forma, que otros tipos de modelos pueden estar presentes.



Gráficamente también se observa que los residuos producidos en la estimación por regresión individual de cada uno de los establecimientos que componen el panel muestran una menor variabilidad que los errores producidos en la estimación EF. Por tanto el contraste y el gráfico nos muestran resultados complementarios y que ambos van en la misma línea.

En conclusión y a la vista de los resultados, parece más apropiado un modelo para cada uno de los formatos y establecimientos (para cada unidad del panel) que un modelo más global como los propuestos. Antes de dar por zanjada esta cuestión vamos a explorar una vía más, la inclusión de variables desfasadas de la variable a explicar cómo variables explicativas. Este planteamiento nos llevara a utilizar modelos dinámicos y será abordado en la siguiente sección.

### 5.4.3.3.- Estimación Dinámica del modelo

El modelo a utilizar es básicamente el mismo que en el caso anterior, la única diferencia es que añadiremos como variables explicativas la variable  $y$  desfasada hasta 4 veces. La razón es que la variable  $y$  está medida semanalmente, de manera que es posible que en dicha variable influyan los valores que tomó la variable hace 1, 2, 3 y 4 semanas antes. En cualquier caso es una hipótesis de partida que intentaremos contrastar posteriormente. Además y tal y como reconocen Hill, Griffiths y Lim, en su libro *Principles of Econometrics* (2012; 365) la inclusión de variables desfasadas, ya sean endógenas o exógenas, ayuda a resolver problemas de autocorrelación presentes en el modelo. El número de desfases a incluir dependerá de la función de autocorrelación parcial de los residuos o de los contrastes de significación aplicados sobre los  $\beta$  asociados a las variables desfasadas, en el sentido de si son significativos esas variables deberán mantenerse en el modelo. Esta es la línea de trabajo que se ha seguido.

La principal diferencia respecto a la estimación anterior se produce en el método de estimación, los estimadores que hemos utilizado anteriormente presentan problemas que es necesario resolver.

El modelo a estimar es de la forma  $y_{it} = \gamma y_{it-1} + \mathbf{x}_{it}'\beta + \alpha_i + u_{it}$ , asumiendo que  $|\gamma| < 1$

Consideremos primeramente la estimación por MCO de  $y_{it}$  sobre  $y_{it-1}$  y  $\mathbf{x}_{it}$ . El término de error  $(\alpha_i + u_{it})$ , está correlacionado con el regresor  $y_{it-1}$  ya que en la ecuación desfasada  $y_{it-1} = \gamma y_{it-2} + \mathbf{x}_{it-1}'\beta + \alpha_i + u_{it-1}$  se muestra la relación entre  $y_{it-1}$  y  $\alpha_i$ ; por lo tanto el estimador resultante será **inconsistente**.

Consideremos ahora el estimador *within* o estimador de Efectos Fijos. Como recordamos, este estimador calcula la regresión de la variable  $(y_{it} - \bar{y}_i)$  sobre las variables explicativas  $(y_{it-1} - \bar{y}_{i-1})$  y  $(\mathbf{x}_{it} - \bar{\mathbf{x}}_i)$ . Esta regresión tiene como término de error a  $(u_{it} - \bar{u}_i)$ . Ahora y gracias a la ecuación  $y_{it} = \gamma y_{it-1} + \mathbf{x}_{it}'\beta + \alpha_i + u_{it}$ ,  $y_{it}$  está



correlacionado con  $u_{it}$  por tanto  $y_{it-1}$  está correlacionado con  $u_{it-1}$  y además con  $\bar{u}_i$ . Esta relación que hemos establecido tiene sus implicaciones ya que el regresor  $(y_{it-1} - \bar{y}_{i-1})$  está correlacionado con el error  $(u_{it} - \bar{u}_i)$ . Por lo tanto la estimación por el estimador de Efectos Fijos produce estimaciones **inconsistentes** de los parámetros.

Esta aportación fue realizada por Nickell (1981). En este artículo Nickell demuestra que el estimador de Efectos Fijos está sesgado por  $O(1/T)$ , y este sesgo no desaparece cuando el número de individuos del panel aumenta, por lo tanto el estimador de Efectos Fijos continúa siendo inconsistente para el caso de  $N$  y  $T$  pequeña. Sin embargo cuando  $T$  es grande, el estimador de Efectos Fijos resulta ser **consistente**. Se han hecho bastantes sugerencias para corregir el sesgo del estimador de Efectos Fijos, quizás la más notable fue realizada por Kiviet (1995).

El estimador de Efectos Aleatorios resulta ser también sesgado en estos modelos dinámicos de datos de panel. La aplicación del estimador de Efectos Aleatorios requiere que realicemos la transformación  $(y_{it} - \theta \bar{y}_i)$  por lo que  $(y_{it-1} - \theta \bar{y}_{i-1})$  estará correlacionado con  $(u_{it} - \theta \bar{u}_i)$ .

En el caso que nos ocupa y gracias a la aportación de Nickell (1981) utilizaremos el estimador de Efectos Fijos ya que  $T$  es bastante grande comparado con los tamaños habituales de otros paneles y  $N$  bastante pequeño.

Los resultados de la estimación por Efectos Fijos son los siguientes:

```
. . xtreg logproporcion L.logproporcion L2.logproporcion L3.logproporcion L4.lo
> gproporcion x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 x13 x14 x15 x16 x17 x18 x19
> x20 x21 x22 x23 x24 x25 x26 x27 x28 x29 x30 x31 x32 x33 x34 x35, fe vce(clust
> er panel)
```

```
Fixed-effects (within) regression                               Number of obs   =    2496
Group variable: panel                                         Number of groups  =     24

R-sq:  within = 0.5418                                         obs per group: min =    104
        between = 0.9861                                       avg              =   104.0
        overall = 0.9561                                       max              =    104

corr(u_i, Xb) = 0.9027                                         F(23,23)         =      .
                                                Prob > F          =      .
```

(Std. Err. adjusted for 24 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
logproporc~n						
L1.	.3893003	.0435284	8.94	0.000	.2992549	.4793456
L2.	.0445325	.0301592	1.48	0.153	-.0178566	.1069217
L3.	.067375	.0192	3.51	0.002	.0276568	.1070932
L4.	.1053179	.0232451	4.53	0.000	.0572317	.1534041
x1	-.0007101	.0013832	-0.51	0.613	-.0035714	.0021512
x2	.6538195	.1611606	4.06	0.000	.3204334	.9872056
x3	.1153924	.0972449	1.19	0.247	-.0857741	.3165589
x4	-.1276088	.1059184	-1.20	0.241	-.3467177	.0915001
x5	-.0004582	.000917	-0.50	0.622	-.0023551	.0014387
x6	-.0009677	.0010019	-0.97	0.344	-.0030404	.001105
x7	-.0006736	.0007008	-0.96	0.346	-.0021233	.0007762
x8	.1788215	.0928909	1.93	0.067	-.013338	.3709811
x9	.0065083	.0019058	3.42	0.002	.0025659	.0104507
x10	.0061323	.0008741	7.02	0.000	.0043241	.0079405
x11	.0044104	.0009578	4.60	0.000	.002429	.0063919
x12	-.0019477	.0006115	-3.18	0.004	-.0032128	-.0006826
x13	.0043188	.0008358	5.17	0.000	.0025898	.0060478
x14	.0004248	.000699	0.61	0.549	-.0010211	.0018708
x15	-.122405	.1157957	-1.06	0.301	-.3619467	.1171367
x16	-.0010854	.0005439	-2.00	0.058	-.0022105	.0000397
x17	.0011201	.0008576	1.31	0.204	-.0006539	.0028941
x18	.0005468	.00043	1.27	0.216	-.0003428	.0014364
x19	-.0002212	.0009963	-0.22	0.826	-.0022822	.0018398
x20	.0013284	.0005122	2.59	0.016	.0002688	.002388
x21	.0000724	.0009484	0.08	0.940	-.0018895	.0020342
x22	.0391669	.0438251	0.89	0.381	-.0514921	.129826
x23	.0010906	.0018903	0.58	0.570	-.0028198	.005001
x24	.0002812	.0006893	0.41	0.687	-.0011447	.0017072
x25	.0007871	.000584	1.35	0.191	-.0004209	.0019952
x26	.0002696	.0003663	0.74	0.469	-.0004881	.0010274
x27	-.0010872	.0003468	-3.14	0.005	-.0018045	-.0003699
x28	.0004961	.0020909	0.24	0.815	-.0038292	.0048214
x29	-.0554621	.0790372	-0.70	0.490	-.218963	.1080387
x30	-.0012626	.0010315	-1.22	0.233	-.0033963	.0008712
x31	.0010258	.0008945	1.15	0.263	-.0008246	.0028761
x32	-.0004622	.0003181	-1.45	0.160	-.0011203	.0001959
x33	-.0023045	.0010011	-2.30	0.031	-.0043754	-.0002336
x34	-.0002919	.0002043	-1.43	0.166	-.0007145	.0001306
x35	-.0482246	.0164951	-2.92	0.008	-.0823473	-.0141019
_cons	-1.586862	.3209359	-4.94	0.000	-2.250768	-.9229554
sigma_u	.38622665					
sigma_e	.19193867					
rho	.80194527	(fraction of variance due to u_i)				

Aunque aparentemente hemos mejorado en el modelo respecto al modelo anterior, el camino a seguir continúa siendo el mismo, ¿es preferible un modelo global para todos los establecimientos y formatos o un modelo para cada establecimiento y

formato? Esta pregunta que ya fue contestada en el modelo anterior debemos contestarla en este nuevo escenario.

Tal y como ya habíamos hecho, se trata de contrastar la hipótesis  $H_1: \beta_1 = \beta_2 = \dots \beta_N$ . Aceptar  $H_1$  significa que el modelo adecuado es  $y_{it} = \alpha_i + \beta' x_{it} + u_{it}$  mientras que rechazarla supone admitir que este modelo no es el adecuado y por tanto, tenemos un modelo para cada individuo.

El estadístico de contraste que se utiliza es: 
$$F_1 = \frac{\frac{S_2 - S_1}{(N-1)K}}{\frac{S_1}{NT - N(K+1)}}; \text{ rechazando la}$$

hipótesis si  $F_1 \geq F_{(N-1)K, NT-N(K+1)}^{(\alpha)}$ .

Para este caso concreto los valores del estadístico son

- $S_2 = 89,63282519$  es la suma de residuos al cuadrado de la regresión por Efectos Fijos.
- $S_1$  es la suma de los residuos al cuadrado. Para calcularlo hacemos una regresión para cada individuo y sumamos los residuos al cuadrado de todas y cada una de las regresiones realizadas (una por individuo). Su valor en este caso es: 52,02903
- $N$  número de individuos que componen el panel y en este caso es 24
- $T$  número de periodos temporales considerados; en este caso hemos perdido periodos porque al utilizar variables desfasadas perdemos 4 periodos, por lo tanto el número de periodos temporales es 104
- $K$  número de variables explicativas consideradas; en este caso han aumentado ya que a las variables originales les hemos añadido las variable desfasadas, por tanto en este caso son 40, incluyendo la constante

Sustituyendo en la expresión del estadístico

$$F_1 = \frac{\frac{89,63282519 - 52,02903}{(24-1) \cdot 40}}{\frac{52,02903}{24 \cdot 104 - 24 \cdot (40+1)}} = 1900,508497$$

Comparándolo con el valor en tablas tenemos:

$$1900,508497 = F_1 \geq F_{(N-1)K, NT-N(K+1)}^{(\alpha)} = F_{(24-1) \cdot 40, 24 \cdot 104 - 24 \cdot (40+1)}^{(5\%)} = F_{920, 1512}^{(5\%)} = 1,10146$$

por lo tanto no se acepta la hipótesis nula.

#### 5.4.3.3.1.- Conclusiones de este modelo.

A la vista de los resultados obtenidos podemos ir estableciendo algunas conclusiones respecto a las hipótesis planteadas en esta tesis.

La primera, la construcción de un modelo que explique el comportamiento de la cuota de mercado en función de las variables de marketing, se ha demostrado como posible. Esto significa que podemos explicar el comportamiento de la cuota de mercado de una marca en función de las acciones comerciales propias y las de sus competidores.

La utilización de la técnica de los datos de panel, en cualquiera de los métodos de estimación utilizados, ha puesto de manifiesto que no existe un único modelo para todos los formatos de la marca. Es decir, los  $\beta$  del modelo son distintos para cada uno de los individuos; por lo que no podemos hablar de un modelo único, sino de varios modelos. El tema es ¿cuántos modelos?, a esta pregunta intentaremos darle respuesta en las secciones siguientes.

Desde el punto de vista de marketing, las conclusiones aparecen, lógicamente, desde otra óptica. El contraste de agrupabilidad utilizado puede traducirse como ¿existe una sola política comercial o bien hay una política comercial para cada formato/establecimiento? Con las conclusiones obtenidas del contraste, estamos en condiciones de afirmar que no existe una única política comercial, pero todavía es pronto para afirmar que existen tantas políticas comerciales como formatos y establecimientos existen. Por lo que sigue iremos trabajando desde esta posición, esto significa que dividiremos la muestra en función de los establecimientos y en función de los formatos, para de esta forma contestar a la pregunta de cuantos modelos o políticas comerciales existen.

#### **5.4.4.- Modelo de datos de panel por establecimientos o por formatos.**

Vamos a considerar dos unidades de agrupación diferentes:

- Los establecimientos, por lo que nos encontraremos con cuatro paneles distintos formados por Hipermercados, Supermercados de 1000 a 2499 m, Supermercados de 400 a 999 m y Supermercados de 100 a 399 m.
- Los formatos, en este caso tendremos 6 paneles distintos, los formados por los formatos Hasta 350 g., De 350 a 499 g, De 500 a 899 g, De 900 a 1200 g y De 3000 a 4999 g.

Tanto en los establecimientos como en los formatos partimos de la hipótesis que los formatos/establecimientos que trabajan siguen el mismo modelo, esto es, los  $\beta$  asociados a ellos son los mismos dentro del establecimiento/formato.

Lógicamente serán distintos a los de los otros establecimientos/formatos tal y como hemos demostrado indirectamente con los contrastes anteriores, pues si bien hemos demostrado heterogeneidad entre los  $\beta$ , no hemos demostrado el origen de esa heterogeneidad, que bien puede partir de 3 fuentes distintas:

4. Heterogeneidad producida por el tipo de establecimiento.
5. Heterogeneidad producida por el formato de producto.
6. Heterogeneidad debida a la interacción tipo de establecimiento y formato de producto.

Es esta última la que se ha puesto de manifiesto en los contrastes de agrupabilidad anteriores, rechazando la existencia de un modelo único frente a un modelo específico para cada formato/establecimiento. A continuación vamos a intentar contrastar la existencia de la heterogeneidad de tipo 1 y 2.

La heterogeneidad producida por el tipo de establecimiento, desde el punto de vista de marketing implica contrastar la hipótesis de que existe un modelo para cada tipo

de establecimiento frente a la no existencia de este modelo. Esta es una cuestión importante, como lo demuestra la existencia de figuras comerciales como el Key Account Manager en los Departamentos Comerciales de las empresas, que vislumbran la posibilidad de un tratamiento personalizado para cada tipo de establecimiento. Será por lo tanto esta hipótesis la que intentaremos poner a prueba.

En el segundo caso, cuando deseamos contrastar la heterogeneidad producida por el formato, desde la óptica de marketing significa dar un tratamiento homogeneizado al formato, independientemente del tipo de establecimiento donde se esté comercializando. Esta ha sido la práctica habitual en los departamentos de marketing, práctica que dada la evolución y concentración de la distribución se supone que ha debido ir cambiando con el paso del tiempo. Esta es la segunda hipótesis que pondremos a prueba.

#### 5.4.4.1.- Contraste de Agrupabilidad por Establecimientos.

Recordemos nuevamente el contraste a utilizar:

El estadístico de contraste que se utiliza es: 
$$F_1 = \frac{\frac{S_2 - S_1}{(N-1)K}}{\frac{S_1}{NT - N(K+1)}}; \text{ rechazando la}$$

hipótesis si  $F_1 \geq F_{(N-1)K, NT-N(K+1)}^{(\alpha)}$ .

Para este caso concreto las variables que forman el estadístico son

- $S_2$  es la suma de residuos al cuadrado de la regresión por Efectos Fijos.
- $S_1$  es la suma de los residuos al cuadrado. Para calcularlo hacemos una regresión para cada individuo y sumamos los residuos al cuadrado de todas y cada una de las regresiones realizadas (una por individuo).

- $N$  número de individuos que componen el panel y en este caso es 6, los seis formatos que consideramos por cada tipo de establecimiento.
- $T$  número de periodos temporales considerados; en este caso hemos perdido periodos porque al utilizar variables desfasadas perdemos 4 periodos, por lo tanto el número de periodos temporales es 104
- $K$  número de variables explicativas consideradas; en este caso han aumentado ya que a las variables originales les hemos añadido las variable desfasadas, por tanto en este caso son 40, incluyendo la constante

Los valores de este contraste para los cuatro tipos de establecimientos se recogen en la tabla siguiente.

	$S_2$	$S_1$	Grados libertad numerador	Grados libertad denominador	Estadístico	P valor
Supermercados de 100 a 399 m	24,03879582 1	15,37940472 4	200	378	1,064166622	0,302663
Supermercados de 400 a 999 m	16,8742183	10,85767369	200	378	4,206311443	1,6346e-033
Supermercados de 1000 a 2499 m	14,22261646	8,393638094	200	378	2,999894872	2,07865e-020
Hipermercados	25,57371613	17,39831753	200	378	0,888103308	0,826399

A la vista de los resultados, las conclusiones que podemos establecer son:

- En los Supermercados de 100 a 399 m y en los Hipermercados podemos aceptar la hipótesis nula, esto es, existe un solo modelo para todos y cada uno de los formatos que se comercializan en estos establecimientos.

Desde el punto de vista de marketing, significa aceptar que para los hipermercados, la figura del Key Account Manager está dando sus frutos, realizando una comercialización del producto homogénea, sin distinción de formatos.

En el caso de los supermercados de 100 a 399 m, en los que también aceptamos la hipótesis, su interpretación en clave de marketing, puede ser



distinta, ya que la figura del Key Account Manager no tiene sentido en estos momentos. Parece más plausible pensar, que el elevado número de establecimientos con estas características, con un elevado número de propietarios, hace que para los fabricantes sea más sencillo tratarles a todos por igual que darles a cada uno un tratamiento específico como sucede en el caso de los hipermercados.

- En los Supermercados de 400 a 999 m y de 1000 a 2499 m. se rechaza la hipótesis nula, esto es, no hay un modelo único para todos, sino que parece haber un modelo para cada uno de los formatos.

Desde la óptica de marketing significa un tratamiento específico para cada formato. Esto parece apuntar a la idea de que en estos tipos de establecimientos es donde se está produciendo uno de los mayores cambios en la distribución española y que de momento resulta imposible un tratamiento homogéneo por formatos. Aun así esta conclusión será revisada un poco más adelante.

#### 5.4.4.2.- Contraste de Agrupabilidad por Formatos.

Volvemos a aplicar el contraste anterior, pero esta vez la unidad de agrupación son los formatos y el panel estará formado por 4 individuos, los establecimientos donde se vende el producto.

En esta ocasión, las variables a utilizar en el estadístico son las mismas que en el caso anterior, la única excepción es:

- $N$  número de individuos que componen el panel y en este caso es 4, los cuatro tipos de establecimiento donde se comercializa cada uno de los formatos.

Los valores de este contraste para los seis formatos se recogen en la tabla siguiente.

	$S_2$	$S_1$	Grados libertad numerador	Grados libertad denominador	Estadístico	P valor
Hasta 350 g.	15,71231129	10,92676755	120	252	0,919726884	0,695486
De 350 a 499 g.	26,70284359	19,88153069	120	252	0,720505746	0,978559
De 500 a 899 g.	3,087548251	1,927066781	120	252	1,26462202	0,062681
De 900 a 1199 g.	5,285753583	3,421138576	120	252	1,144557996	0,188145
De 1200 a 1499 g.	10,24407429	6,95286744	120	252	0,994055252	0,50773
De 3000 a 4999 g.	15,51805041	8,919663005	120	252	1,553490703	0,00194427

A la luz de los resultados, podemos afirmar que:

- Existe un comportamiento homogéneo en el tratamiento de los formatos respecto a los establecimientos.
- Nos encontramos con 5 modelos distintos, uno para cada formato, en los formatos: Hasta 350 g., De 350 a 499 g., De 500 a 899 g., De 900 a 1199 g. y de 1200 a 1499 g. De tal manera que el comportamiento de los distintos tipos de establecimientos dentro del formato es homogéneo.
- El único formato donde se rechaza la hipótesis es en el formato de 3000 a 4999 g. La razón desde el punto de vista económico, que no estadístico, es

que este formato tiene una presencia muy importante en los hipermercados, y son precisamente estos establecimientos los que tienen un comportamiento homogéneo como veíamos en el contraste anterior.

- Por lo tanto estimaremos 6 modelos, 5 modelos que hacen referencia a los formatos (Hasta 350 g., De 350 a 499 g., De 500 a 899 g., De 900 a 1199 g. y de 1200 a 1499 g) y un modelo específico para los hipermercados

Tras estos contrastes podemos afirmar que el comportamiento de la empresa refleja dos dimensiones distintas pero a la vez que se superponen.

Una dimensión vertical, que proporciona un tratamiento homogéneo por formato independientemente del establecimiento y una dimensión horizontal con un tratamiento homogéneo por establecimiento y no por formato. Esta última dimensión resulta especialmente importante en los hipermercados.

Esta dimensión refleja el comportamiento que han realizado las compañías de gran consumo dando un tratamiento personalizado por clientes, y este tratamiento ha comenzado por los hipermercados que son los que tienen un mayor poder para “exigirlo”.

En el futuro y con la concentración de la distribución que se está produciendo en España, especialmente en el supermercado grande y mediano, es posible que este tratamiento se extienda a otros tipos de establecimientos; cosa que hasta el momento no sucede.

Como una imagen, vale más que mil palabras, la siguiente tabla obtenida del Boletín Económico del ICE (2011) nos muestra la evolución y la importancia del supermercado, en la distribución alimentaria en España.

CUADRO 6.6						
EVOLUCIÓN DE LAS VENTAS DE ALIMENTACIÓN ENVASADA POR FORMATOS (Porcentaje)						
	1994	1998	2002	2006	2009	2010
Tradicional	13	9	5,9	4,4	5,1	5,1
Autoservicio hasta 100m2.	12	8,8	6,4	4,8	3,9	3,9
Supermercado pequeño 100-399 m2	9	20,9	19,8	16,4	15,3	15
Supermercado mediano 400-999 m2	15	16,7	21,9	20,8	19,9	19,9
Supermercado grande 1000-2499 m2.	10	12,9	20,3	31,2	36,7	38
Hipermercado 2500 m2 o más	31	31,6	25,7	22,5	19,1	18,2
.....						
TOTAL FORMATOS	100	100	100	100	100	100
Fuente: AC Nielsen.						

## 6.- ¿QUÉ VARIABLES INFLUYEN EN LOS MODELOS?

En el capítulo anterior determinamos el método de estimación a utilizar y la agrupación de los individuos según la homogeneidad de sus comportamientos. En este capítulo se intentará dar respuesta a las preguntas:

- ¿Qué variables son las que deben ser incluidas en cada uno de los modelos de datos de panel que hemos estimado?
- ¿Son las mismas variables en cada modelo o por el contrario cambian según los modelos?
- ¿La influencia de esas variables es idéntica en cada una de las unidades de panel consideradas?

La respuesta tiene que ser tanto desde el punto de vista estadístico como económico o de marketing.

La forma de proceder será, estimar los modelos, eliminar aquellas variables que no sean significativas y quedarnos con los modelos más manejables y sencillos.

Esto lo haremos para los 6 modelos considerados, uno para cada formato, menos el formato de 3000 a 4999 g., que será sustituido por el modelo para hipermercados.

## 6.1.- ESTIMACIÓN DEL MODELO DINÁMICO DE EFECTOS INDIVIDUALES ESPECÍFICOS PARA EL FORMATO HASTA 350 G.

El modelo desde el que partimos es el estimado en el capítulo anterior:

```
. xtreg logproporcion_ retardo1logproporcion_ retardo2logproporcion_ retardo3logproporcion_ retardo4logproporcion_ x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 x13 x14 x15 x16 x17 x18 x19 x20 x21 x22 x23 x24 x25 x26 x27 x28 x29 x30 x31 x32 x33 x34 x35, fe vce(cluster panel)
note: x9 omitted because of collinearity
note: x12 omitted because of collinearity
note: x14 omitted because of collinearity
note: x30 omitted because of collinearity
note: x32 omitted because of collinearity
note: x33 omitted because of collinearity
note: x35 omitted because of collinearity
```

```
Fixed-effects (within) regression               Number of obs   =       416
Group variable: panel                          Number of groups =         4

R-sq:  within = 0.3420                        Obs per group:  min =       104
          between = 0.7816                      avg   =      104.0
          overall  = 0.7064                      max   =       104
```

```
corr(u_i, xb) = -0.2107                      F(3,3) = .
                                          Prob > F = .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.2520993	.107392	2.35	0.101	-.08967	.5938686
retardo2lo~n	.1026098	.033114	3.10	0.053	-.0027736	.2079932
retardo3lo~n	.1232988	.0540034	2.28	0.107	-.0485642	.2951618
retardo4lo~n	.0140177	.0166398	0.84	0.461	-.0389376	.066973
x1	-.006635	.0016201	-4.10	0.026	-.0117909	-.0014791
x2	.5540341	.8435442	0.66	0.558	-2.1305	3.238568
x3	-.2429127	.3244404	-0.75	0.508	-1.275427	.7896015
x4	-.0322772	.0647725	-0.50	0.653	-.2384123	.173858
x5	-.0056881	.0023738	-2.40	0.096	-.0132427	.0018664
x6	-.0010715	.0019225	-0.56	0.616	-.0071899	.0050469
x7	.0032351	.0020681	1.56	0.216	-.0033464	.0098167
x8	.2273756	.1288563	1.76	0.176	-.1827027	.6374538
x9	(omitted)					
x10	.0116521	.0084616	1.38	0.262	-.0152766	.0385807
x11	.0015596	.0068431	0.23	0.834	-.0202181	.0233373
x12	(omitted)					
x13	.0042685	.0012755	3.35	0.044	.0002093	.0083277
x14	(omitted)					
x15	-.5307098	.5127106	-1.04	0.377	-2.162384	1.100964
x16	.0019284	.0084811	0.23	0.835	-.0250621	.0289189
x17	-.0175846	.0016461	-10.68	0.002	-.0228233	-.012346
x18	-.0011803	.0016348	-0.72	0.522	-.0063829	.0040224
x19	-.0007198	.0008371	-0.86	0.453	-.0033838	.0019442
x20	.0003532	.0012308	0.29	0.793	-.0035637	.0042702
x21	.0016762	.0016501	1.02	0.385	-.0035753	.0069277
x22	.1375337	.1007431	1.37	0.266	-.1830758	.4581432
x23	-.0200622	.0127652	-1.57	0.214	-.0606866	.0205623
x24	.0021565	.0025712	0.84	0.463	-.0060262	.0103391
x25	-.0008555	.0038323	-0.22	0.838	-.0130517	.0113406
x26	.0027213	.0050935	0.53	0.630	-.0134884	.018931
x27	-.0016859	.0018702	-0.90	0.434	-.0076379	.004266
x28	-.0015248	.0280582	-0.05	0.960	-.0908184	.0877688
x29	-.1317541	.5248158	-0.25	0.818	-1.801952	1.538444
x30	(omitted)					
x31	-.0055267	.0012454	-4.44	0.021	-.0094901	-.0015632
x32	(omitted)					
x33	(omitted)					
x34	-.0002443	.0003154	-0.77	0.495	-.0012479	.0007594
x35	(omitted)					
_cons	-1.55674	.4173535	-3.73	0.034	-2.884945	-.2285352
sigma_u	.29306148					
sigma_e	.20334254					
rho	.6750205	(fraction of variance due to u_i)				

Resulta obvio que este modelo tiene un conjunto muy extenso de variables que es preciso depurar. Esto es lo que se ha hecho hasta llegar al modelo que se muestra a continuación.

```
. xtreg logproporcion retardo1logproporcion retardo2logproporcion_ x7 x13 x31,
> fe vce(cluster panel )
```

```
Fixed-effects (within) regression      Number of obs   =      424
Group variable: panel                 Number of groups =        4

R-sq:  within = 0.2856                obs per group: min =      106
      between = 0.9980                  avg =      106.0
      overall  = 0.8265                  max =      106

corr(u_i, Xb) = 0.8844                F(3,3)          =      .
                                      Prob > F          =      .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.3137212	.0640039	4.90	0.016	.1100322	.5174102
retardo2lo~n	.1831895	.0730228	2.51	0.087	-.0492015	.4155805
x7	.0040586	.0007233	5.61	0.011	.0017566	.0063605
x13	.0056823	.0005033	11.29	0.001	.0040805	.007284
x31	-.0051854	.001506	-3.44	0.041	-.0099783	-.0003926
_cons	-2.410417	.1378165	-17.49	0.000	-2.84901	-1.971823
sigma_u	.3255191					
sigma_e	.20485778					
rho	.7163058	(fraction of variance due to u_i)				

### 6.1.1.- Variables explicativas.

- **Retardo1 y Retardo2.** Recogen los valores de la variable explicativa desfasados uno y dos periodos, el signo asociado a estas variables es positivo, por lo tanto implica que una evolución positiva en la cuota de mercado supone que el efecto se mantiene durante dos semanas.
- **Distribución Ponderada Marcas de la Distribución:**  $x_7$  Es un índice que nos informa sobre la distribución (% de establecimientos que trabajan el producto) y la calidad de esta distribución (importancia que tienen en las ventas los establecimientos donde está presente la marca). El signo asociado al parámetro  $\beta$  tiene signo positivo.

Desde el punto de vista económico, su interpretación sería: una mayor calidad de la distribución afecta positivamente tanto a la marca de la distribución como a la marca competidora, en este caso la marca 1. Cuando se está presente en mejores establecimientos mayores son las ventas de todas las marcas, incluso las competidoras.

- **Cobertura reducción temporal de precio de la Marca 1:**  $x_{13}$  Cuanto mayor es el número de establecimientos y su importancia que realizan la promoción de precio de la Marca 1 mayores tienen que ser las ventas de la Marca 1.

El signo esperado asociado a esta variable es positivo; esto es, coberturas elevadas en este tipo de promoción deberían estar vinculadas a cuotas de mercado elevadas. Y eso es lo que sucede en este caso, por lo tanto, la realización de este tipo de acciones tiene un efecto positivo sobre la cuota de mercado.

- **Cobertura Promoción Solo Exposición Especial Marcas Distribución:**  $x_{31}$   
En este caso medimos la influencia que tienen las acciones de exposición especial realizadas por las Marcas de la Distribución sobre la marca considerada. El signo es negativo indicando un mayor esfuerzo promocional de las marcas de la distribución producirán un deterioro de la cuota de mercado de la marca 1.

Una forma alternativa de comprobar si efectivamente las variables no incluidas son no significativas consiste en un contraste conjunto sobre este conjunto de variables. Se trata de un contraste de tipo Wald y siguiendo de nuevo a Carter Hill, Griffiths y Lim, (2012, pág. 225), es un contraste donde se comparan la suma de residuos al cuadrado del modelo sin restringir con la suma de residuos al cuadrado del modelo restringido.

En este contraste, la hipótesis nula es que las variables, conjuntamente consideradas, son 0, frente a la alternativa que alguna o algunas de ellas no son 0.



Por lo tanto, estamos comparando los resultados de dos regresiones: una primera regresión donde están incluidas todas las variables (modelo sin restringir) con otra regresión donde no están incluidas todas las variables (modelo restringido). Si los resultados de ambas regresiones, medidos por la suma de los residuos al cuadrado, son parecidos, esto parece decirnos que la hipótesis nula es cierta, ya que eliminadas estas variables los resultados del modelo son los mismos.

Los resultados del contraste son los siguientes.

```

. testparm retardo3logproporcion_ retardo4logproporcion_ x1 x2 x3 x4 x5 x6 x8 x
> 9 x10 x11 x12 x14 x15 x16 x17 x18 x19 x20 x21 x22 x23 x24 x25 x26 x27 x28 x29
> x30 x32 x33 x34 x35

( 1) retardo3logproporcion_ = 0
( 2) retardo4logproporcion_ = 0
( 3) x1 = 0
( 4) x2 = 0
( 5) x3 = 0
( 6) x4 = 0
( 7) x5 = 0
( 8) x6 = 0
( 9) x8 = 0
(10) x10 = 0
(11) x11 = 0
(12) x15 = 0
(13) x16 = 0
(14) x17 = 0
(15) x18 = 0
(16) x19 = 0
(17) x20 = 0
(18) x21 = 0
(19) x22 = 0
(20) x23 = 0
(21) x24 = 0
(22) x25 = 0
(23) x26 = 0
(24) x27 = 0
(25) x28 = 0
(26) x29 = 0
(27) x34 = 0
Constraint 1 dropped
Constraint 2 dropped
Constraint 3 dropped
Constraint 6 dropped
Constraint 7 dropped
Constraint 8 dropped
Constraint 9 dropped
Constraint 10 dropped
Constraint 11 dropped
Constraint 12 dropped
Constraint 13 dropped
Constraint 14 dropped
Constraint 15 dropped
Constraint 16 dropped
Constraint 17 dropped
Constraint 18 dropped
Constraint 19 dropped
Constraint 20 dropped
Constraint 21 dropped
Constraint 22 dropped
Constraint 23 dropped
Constraint 24 dropped
Constraint 25 dropped
Constraint 27 dropped

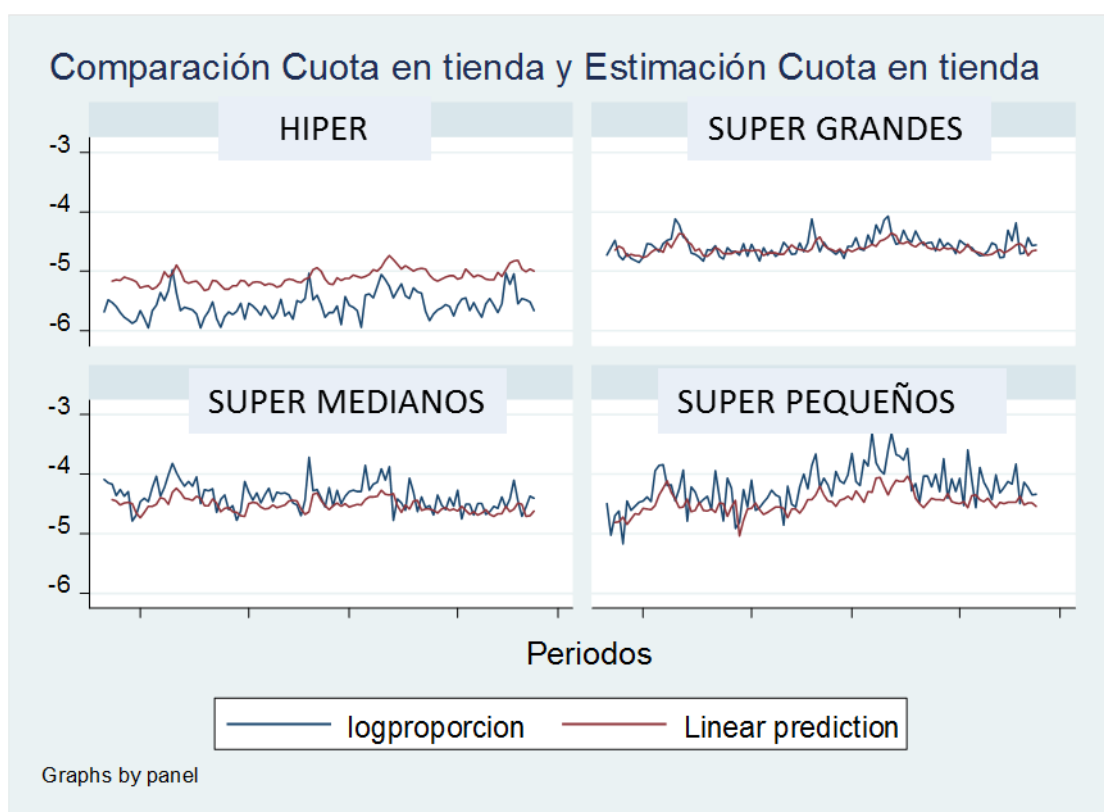
F( 3, 3) = 0.56
Prob > F = 0.6785

```

Por tanto, a la vista de los resultados del contraste, aceptamos la hipótesis nula; esto es, las variables en conjunto consideradas no son significativas y pueden eliminarse del modelo, obteniéndose un modelo más sencillo con muchas menos variables incluidas en él.

Por último y para ilustrar las diferencias existentes entre los distintos individuos que componen el panel (en este caso establecimientos), en el gráfico se muestra la comparación entre la variable  $y$  real y la variable  $y$  estimada. Aunque en los contrastes de agrupabilidad realizados en el capítulo anterior, aceptábamos la utilización de datos de panel, puede verse que aun así existe heterogeneidad en cada uno de los paneles.

Esta heterogeneidad es más pronunciada en el panel 1 ya que la distancia entre la variable dependiente real y la estimada es mayor. Este panel es precisamente, Formato hasta 350 g. en Hipermercados. Y son los Hipermercados los únicos establecimientos donde hemos propuesto un modelo específico para ellos; por lo tanto parece lógico este comportamiento que reflejan los datos.



Por otra parte la existencia de heterogeneidad que reflejan los gráficos, se ve también recogida en el coeficiente de determinación mostrado por el modelo.

En la estimación nos muestra la información siguiente:  $R^2$  global 0,8265,  $R^2$  entre los modelos 0,9980 y  $R^2$  dentro del modelo 0,2856. Este último valor es bajo y debe

interpretarse como que en la estimación de cada uno de los modelos por separado, las variables utilizadas y el peso de cada una de esas variables podrían ser distintas, aunque si bien es cierto que globalmente son las mismas.

Lógicamente el siguiente paso será contrastar esta última afirmación; es decir, comprobar si todos los establecimientos y las variables asociadas a ellos tienen el mismo efecto sobre la variable a explicar. En el contraste de agrupabilidad hemos establecido que elementos del análisis pueden agruparse, pero ¿influyen todos de la misma forma sobre variable a explicar?

Para determinar la influencia de cada uno utilizaremos un contraste que incluya una o varias variables ficticias, e incluso la interacción de la variable ficticia con las variables explicativas utilizadas.

La metodología de este contraste en econometría con datos de sección cruzado o con datos temporales puede encontrarse en el libro de Carter, Griffiths y Lim (2012; pág. 170 y siguientes); pero la aplicación a datos de panel ha sido sugerida por la profesora Rocío Sánchez del Departamento de Economía Cuantitativa de la Universidad Autónoma de Madrid.

En primer lugar formulamos un nuevo modelo, en este nuevo modelo introducimos una variable ficticia. La variable ficticia la denotamos como,  $D_i$   $i = 0,1,2,3$  que hace referencia al tipo de establecimiento, los valores de esta variables serán 1 cuando el establecimiento es el considerado y 0 en el resto. En nuestro caso la categoría de referencia será Supermercados Pequeños.

La formulación del contraste, para un modelo con dos variables explicativas y 4 tipos de establecimientos es:

$$x = \{x_1 x_2\}$$

$$i = 0,1,2,3$$

$$D_i \text{ 1 si tipo } = i; 0 \text{ resto}$$

$$y_{it} = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_3 + \beta_4 x_{1it} + \beta_5 x_{2it} + \beta_6 D_1 x_{1it} + \beta_7 D_1 x_{2it} + \beta_8 D_2 x_{1it} + \beta_9 D_2 x_{2it} + \beta_{10} D_3 x_{1it} + \beta_{11} D_3 x_{2it} + u_{it}$$

Las hipótesis a contrastar con este modelo

$$H_0 : \beta_6 = 0; \beta_7 = 0; \beta_8 = 0; \beta_9 = 0; \beta_{10} = 0; \beta_{11} = 0$$

Aceptar esta hipótesis supone reconocer que la influencia del tipo de establecimiento es la misma siempre. No aceptarla significa reconocer que alguna combinación de variable explicativa y tipo de establecimiento produce un efecto distinto (mayor o menor) sobre la variable a explicar. Por ejemplo, en el modelo planteado podría significar que la variable exposición especial tiene un efecto distinto en los Hipermercados que en los Supermercados Medianos.

El efecto de la variable explicativa  $x_1$  sobre  $y$  se determina de la siguiente forma:

$\beta_4 \Rightarrow$  Efecto de  $x_1$  sobre  $y$  en establecimiento de referencia

$\beta_4 + \beta_6 \Rightarrow$  Efecto de  $x_1$  sobre  $y$  en establecimiento 1

$\beta_4 + \beta_8 \Rightarrow$  Efecto de  $x_1$  sobre  $y$  en establecimiento 2

$\beta_4 + \beta_{10} \Rightarrow$  Efecto de  $x_1$  sobre  $y$  en establecimiento 3

Para realizar este contraste en lugar de estimar el modelo por Efectos Fijos utilizaremos una propiedad que tiene este estimador y es que podemos obtener el mismo resultado realizando una regresión sobre la variable original y añadiendo en el modelo tantas variables ficticias como tipos de establecimientos menos uno. Es por esta razón por la que el estimador de Efectos Fijos es conocido también como el estimador de Variable Ficticia. Lógicamente utilizaremos como variables explicativas todas las que han salido significativas en el modelo.

La aplicación concreta de este método está reflejada en la salida que vemos a continuación:

```
. regress logproporcion retardo1logproporcion retardo2logproporcion_ x7 x13 x31 su
> perm superg hiper r1h r1sg r1sm r2h r2sg r2sm intx7hiper intx7sg intx7sm intx13h
> iper intx13sg intx13sm intx31hiper intx31sg intx31sm, vce(cluster panel)
note: intx31sg omitted because of collinearity
note: intx31sm omitted because of collinearity
```

Linear regression

Number of obs = 424  
F( 0, 3) = .  
Prob > F = .  
R-squared = 0.8817  
Root MSE = .20489

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.2212802	3.08e-12	7.2e+10	0.000	.2212802	.2212802
retardo2lo~n	.2728361	1.72e-13	1.6e+12	0.000	.2728361	.2728361
x7	.0566944	1.96e-13	2.9e+11	0.000	.0566944	.0566944
x13	.0066769	4.70e-14	1.4e+11	0.000	.0066769	.0066769
x31	-.0076975	3.55e-14	-2.2e+11	0.000	-.0076975	-.0076975
superm	-.4849988	1.44e-11	-3.4e+10	0.000	-.4849988	-.4849988
superg	-.1827797	2.08e-11	-8.8e+09	0.000	-.1827797	-.1827797
hiper	-.4280796	5.37e-11	-8.0e+09	0.000	-.4280796	-.4280796
r1h	.3011074	6.61e-12	4.6e+10	0.000	.3011074	.3011074
r1sg	.2171905	7.12e-12	3.0e+10	0.000	.2171905	.2171905
r1sm	.0657192	3.08e-12	2.1e+10	0.000	.0657192	.0657192
r2h	-.2647209	3.32e-12	-8.0e+10	0.000	-.2647209	-.2647209
r2sg	-.2219829	3.24e-12	-6.9e+10	0.000	-.2219829	-.2219829
r2sm	-.1695786	6.37e-13	-2.7e+11	0.000	-.1695786	-.1695786
intx7hiper	-.0535149	2.05e-13	-2.6e+11	0.000	-.0535149	-.0535149
intx7sg	-.0480976	2.03e-13	-2.4e+11	0.000	-.0480976	-.0480976
intx7sm	-.0513288	1.96e-13	-2.6e+11	0.000	-.0513288	-.0513288
intx13hiper	-.0010142	4.70e-14	-2.2e+10	0.000	-.0010142	-.0010142
intx13sg	-.0020622	4.76e-14	-4.3e+10	0.000	-.0020622	-.0020622
intx13sm	.0010878	4.77e-14	2.3e+10	0.000	.0010878	.0010878
intx31hiper	.0061955	4.56e-14	1.4e+11	0.000	.0061955	.0061955
intx31sg	(omitted)					
intx31sm	(omitted)					
_cons	-2.255419	1.43e-11	-1.6e+11	0.000	-2.255419	-2.255419

Las variables explicativas son:

1. **retardo1 y retardo2.** Recogen los valores de la variable explicativa desfasados uno y dos periodos.
2. **Distribución Ponderada Marcas de la Distribución:**  $x_7$
3. **Cobertura reducción temporal de precio de la Marca 1:**  $x_{13}$
4. **Cobertura Promoción Solo Exposición Especial Marcas Distribución:**  $x_{31}$
5. **Superm.** Toma el valor 1 cuando se trata de supermercados medianos, en los demás casos 0.

- 6. *Superg.*** Toma el valor 1 cuando se trata de supermercados grandes, en los demás casos 0.
- 7. *Hiper.*** Toma el valor 1 cuando se trata de Hipermercado, en los demás casos 0. Supermercados pequeños es la categoría de referencia.
- 8. *r1h.*** Mide la interacción entre la variable explicativa desfasada un periodo y los hipermercados. Mide el efecto diferencial de esta variable en los hipermercados.
- 9. *r1sg.*** Mide el efecto diferencial de la variable  $r1$  en los supermercados grandes.
- 10. *r1sm.*** Mide el efecto diferencial de la variable  $r1$  en los supermercados medianos.
- 11. *r2h.*** Mide la interacción entre la variable explicativa desfasada dos periodos en los hipermercados.
- 12. *r2sg.*** Mide el efecto diferencial de la variable  $r2$  en los supermercados grandes.
- 13. *r2sm.*** Mide el efecto diferencial de la variable  $r2$  en los supermercados medianos.
- 14. *Intx7hiper.*** Esta variable mide la interacción entre la variable Distribución ponderada Marcas de la Distribución y los hipermercados. Aquí es donde se trata de medir el efecto diferencial de la variable  $x_7$  en los hipermercados. De tal manera que, si en el contraste que vamos a realizar aceptamos, que el parámetro asociado a esta variable es cero supone admitir que esta variable produce el mismo efecto sobre la variable explicativa, en los Supermercados pequeños que en los Hipermercados. Esta argumentación será la misma para las variables que vienen a continuación.

**15. *Intx7sg*.** Mide el efecto diferencial de la variable  $x_7$  en los supermercados grandes.

**16. *Intx7sm*.** Es el efecto diferencial de la variable  $x_7$  en los supermercados medianos

**17. *Intx13hiper*.** Mide el efecto diferencial de la variable  $x_{13}$  en los hipermercados

**18. *Intx13sg*.** Es el efecto diferencial de la variable  $x_{13}$  en los supermercados grandes.

**19. *Intx13sm*.** Es el efecto diferencial de la variable  $x_{13}$  en los supermercados medianos.

**20. *Intx31hiper*.** Mide el efecto diferencial de la variable  $x_{31}$

**21. *Intx31sg*.** Es el efecto diferencial de la variable  $x_{31}$  en los supermercados grandes.

**22. *Intx31sm*.** Es el efecto diferencial de la variable  $x_{31}$  en los supermercados medianos.

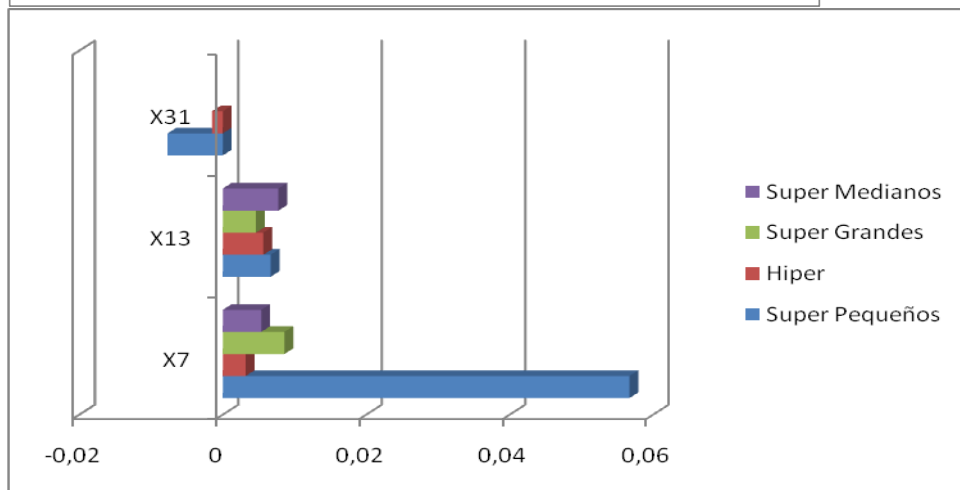
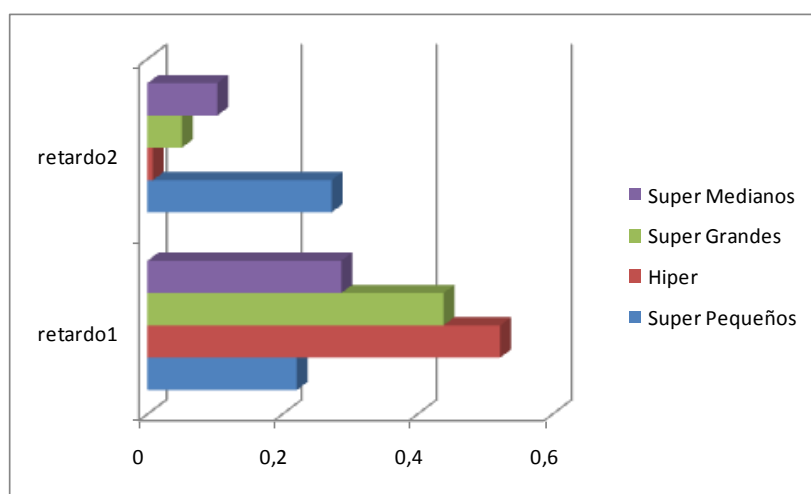
Para comprobar si efectivamente las nuevas variables introducidas son estadísticamente significativas nos fijamos en los estadísticos t y en el P valor asociado. Como resulta que el p valor de todas las estimaciones es 0, rechazamos la hipótesis nula y podemos concluir que existe un efecto diferencial de estas variables según estén referidas a los 3 tipos de establecimientos en comparación con los supermercados pequeños.

Efectos de las variables según los tipos de establecimientos:



	<i>Super Pequeños</i>	<i>Hiper</i>	<i>Super Grandes</i>	<i>Super Medianos</i>
<i>retardo1</i>	0,2212802	0,5223876	0,4384707	0,2869994
<i>retardo2</i>	0,2728361	0,0081152	0,0508532	0,1032575
$X_7$	0,0566944	0,0031795	0,0085968	0,0053656
$X_{13}$	0,0066769	0,0056627	0,0046147	0,0077647
$X_{31}$	-0,0076975	-0,001502		
<i>Constante</i>	-2,255419	-2,6834986	-2,4381987	-2,7404178

Efectivamente los efectos de las variables no son los mismos según del tipo de establecimiento de que se trata.



El último paso será comprobar que la forma funcional utilizada es la adecuada, para ello vamos a utilizar el contraste RESET (Regression Specification Error Test). Este contraste, como establecen Carter Hill, Griffiths y Lim, (2012; pág. 151), está diseñado para detectar variables omitidas y formas funcionales incorrectas. Su funcionamiento es el siguiente.

En un primer paso se estima el modelo y se obtienen los valores ajustados del

$$\begin{aligned} \text{modelo} \quad y_i &= \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + u_i \\ \hat{y}_i &= b_1 + b_2 x_{i2} + b_3 x_{i3} \end{aligned}$$

En un segundo momento, se considera los dos modelos artificiales siguientes:

$$\begin{aligned} y_i &= \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \gamma_1 \hat{y}_i^2 + u_i \\ y_i &= \beta_1 + \beta_2 x_{i2} + \beta_3 x_{i3} + \gamma_1 \hat{y}_i^2 + \gamma_2 \hat{y}_i^3 + u_i \end{aligned}$$

Un contraste de error de especificación es testar si  $H_o: \gamma_1 = 0$  frente a la hipótesis alternativa  $H_1: \gamma_1 \neq 0$  en el primer modelo. En el segundo modelo se contrasta la hipótesis:  $H_o: \gamma_1 = \gamma_2 = 0$  frente a  $H_1: \gamma_1 \neq 0$  y/o  $\gamma_2 \neq 0$ . En el primer caso podemos utilizar un contraste  $t$  o  $F$ . En el segundo caso utilizaremos un test  $F$ .

Rechazar la hipótesis nula implica que el modelo original no es adecuado y puede ser mejorado. No aceptar la hipótesis nula significa que el test no es capaz de detectar infraparametrización. Recordemos que  $\hat{y}_i^2$  y  $\hat{y}_i^3$  son funciones polinomiales de  $x_{i2}$  y  $x_{i3}$  (como por ejemplo  $x_{i2}^2, x_{i3}^3, x_{i2}x_{i3}, x_{i2}x_{i3}^2$ ) Estos polinomios pueden aproximar muchos tipos de formas funcionales, si la forma funcional original no es correcta la aproximación polinomial que incluya  $\hat{y}_i^2$  y  $\hat{y}_i^3$  puede mejorar significativamente el ajuste del modelo. No sólo eso, si hemos omitido variables, y esas variables están correlacionadas con  $x_{i2}$  y  $x_{i3}$ , entonces también estarán correlacionadas con términos como  $\hat{y}_i^2$  y  $\hat{y}_i^3$  y alguno de sus efectos pueden ser recogidos al incluir los términos  $\hat{y}_i^2$  y  $\hat{y}_i^3$ .

La filosofía general del test es: si podemos mejorar significativamente el modelo aumentando artificialmente su poder de predicción, entonces el modelo original no es adecuado.

Este contraste es el que se ha seguido, y como se decía en el capítulo 3, se ha aplicado a datos de panel; los resultados obtenidos son los siguientes.

```
. xtreg logproporcion retardo1logproporcion retardo2logproporcion_ x7 x13 x31
> pred2, fe vce(cluster panel)
```

```
Fixed-effects (within) regression               Number of obs   =       424
Group variable: panel                          Number of groups =        4

R-sq:  within = 0.2858                        Obs per group:  min =       106
          between = 0.9984                      avg   =      106.0
          overall = 0.8207                      max   =       106

corr(u_i, Xb) = 0.8783                        F(3,3)          =        .
                                          Prob > F         =        .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.4947534	.2212405	2.24	0.111	-.2093325	1.198839
retardo2lo~n	.2890691	.1898009	1.52	0.225	-.3149621	.8931003
x7	.0067915	.0046214	1.47	0.238	-.0079159	.0214989
x13	.0089275	.0048139	1.85	0.161	-.0063924	.0242475
x31	-.0083869	.005834	-1.44	0.246	-.0269532	.0101794
pred2	.063228	.0907751	0.70	0.536	-.2256588	.3521148
_cons	-2.492482	.2373549	-10.50	0.002	-3.247851	-1.737113
sigma_u	.3352214					
sigma_e	.2050725					
rho	.72767441	(fraction of variance due to u_i)				

```
. testparm pred2
```

```
( 1)  pred2 = 0
```

```
      F( 1, 3) = 0.49
      Prob > F = 0.5362
```

Puede comprobarse que la variable pred2, que recogen la variable estimada al cuadrado, resultan ser no significativas.

```
. xtreg logproporcion retardo1logproporcion retardo2logproporcion_ x7 x13 x31
> pred2 pred3, fe vce(cluster panel )
```

```
Fixed-effects (within) regression      Number of obs   =      424
Group variable: panel                 Number of groups =        4

R-sq:  within = 0.2859                obs per group: min =      106
      between = 0.9985                avg =      106.0
      overall  = 0.8199                max =      106
```

```
corr(u_i, Xb) = 0.8775                F(3,3) = .
                                      Prob > F = .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	2.013815	5.09821	0.40	0.719	-14.21096	18.23859
retardo2lo~n	1.175973	2.960966	0.40	0.718	-8.247144	10.59909
x7	.026373	.0664577	0.40	0.718	-.185125	.2378709
x13	.036392	.0927435	0.39	0.721	-.2587592	.3315432
x31	-.0335534	.0849175	-0.40	0.719	-.3037988	.2366921
pred2	1.094886	3.421363	0.32	0.770	-9.793418	11.98319
pred3	.0730011	.2387587	0.31	0.780	-.6868355	.8328378
_cons	-6.615253	13.67958	-0.48	0.662	-50.14979	36.91929
sigma_u	.33653524					
sigma_e	.20531294					
rho	.72875895	(fraction of variance due to u_i)				

```
. testparm pred2 pred3
```

```
( 1) pred2 = 0
( 2) pred3 = 0
```

```
F( 2, 3) = 0.32
Prob > F = 0.7499
```

Aquí las variables pred2 y pred3 (las variables estimadas al cuadrado y al cubo) son nuevamente no significativas; por lo tanto estamos ante una situación donde no hay omisión de variables ni forma funcional incorrecta.

Por tanto el modelo estimado como definitivo es el que se presenta en el cuadro siguiente:

```
. xtreg logproporcion retardo1logproporcion retardo2logproporcion_ x7 x13 x31,
> fe vce(cluster panel )
```

```
Fixed-effects (within) regression      Number of obs   =      424
Group variable: panel                 Number of groups =        4

R-sq:  within = 0.2856                obs per group: min =      106
      between = 0.9980                avg =      106.0
      overall  = 0.8265                max =      106
```

```
corr(u_i, Xb) = 0.8844                F(3,3) = .
                                      Prob > F = .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.3137212	.0640039	4.90	0.016	.1100322	.5174102
retardo2lo~n	.1831895	.0730228	2.51	0.087	-.0492015	.4155805
x7	.0040586	.0007233	5.61	0.011	.0017566	.0063605
x13	.0056823	.0005033	11.29	0.001	.0040805	.007284
x31	-.0051854	.001506	-3.44	0.041	-.0099783	-.0003926
_cons	-2.410417	.1378165	-17.49	0.000	-2.84901	-1.971823
sigma_u	.3255191					
sigma_e	.20485778					
rho	.7163058	(fraction of variance due to u_i)				

## 6.2.- ESTIMACIÓN DEL MODELO DINÁMICO DE EFECTOS INDIVIDUALES ESPECÍFICOS FORMATO DE 350 A 499 g.

Este es el modelo de partida en el formato de 350 a 499 g. a partir de este modelo iremos eliminando las variables que no aportan ninguna información.

```
. xtreg logproporcion L.logproporcion L2.logproporcion L3.logproporcion L4.logp
> roporcion x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 x13 x14 x15 x16 x17 x18 x19 x
> 20 x21 x22 x23 x24 x25 x26 x27 x28 x29 x30 x31 x32 x33 x34 x35, fe vce(cluste
> r panel )
note: x35 omitted because of collinearity
```

```
Fixed-effects (within) regression               Number of obs   =       416
Group variable: panel                          Number of groups =         4

R-sq:  within = 0.3115                        Obs per group:  min =       104
          between = 0.8362                      avg   =      104.0
          overall = 0.5052                      max   =       104

corr(u_i, Xb) = -0.0734                       F(3,3)          =         .
                                          Prob > F         =         .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
logproporc~n						
L1.	.3089082	.0378258	8.17	0.004	.1885297	.4292866
L2.	.0432001	.0557651	0.77	0.495	-.1342695	.2206696
L3.	.0570954	.0350611	1.63	0.202	-.0544848	.1686755
L4.	.1548991	.0564755	2.74	0.071	-.0248311	.3346292
x1	-.0103132	.0057039	-1.81	0.168	-.0284656	.0078393
x2	-2.282451	.1689684	-13.51	0.001	-2.820184	-1.744718
x3	.1181029	.1264945	0.93	0.419	-.2844592	.520665
x4	.0486619	.4998839	0.10	0.929	-1.542192	1.639515
x5	.011589	.0051627	2.24	0.110	-.0048408	.0280189
x6	-.0000509	.0023016	-0.02	0.984	-.0073756	.0072738
x7	-.00567	.0016479	-3.44	0.041	-.0109143	-.0004257
x8	-.7100222	.4195847	-1.69	0.189	-2.045328	.6252836
x9	.041918	.0059651	7.03	0.006	.0229342	.0609017
x10	.0144204	.0026684	5.40	0.012	.0059285	.0229123
x11	.016787	.0030624	5.48	0.012	.0070411	.0265329
x12	.032874	.0039238	8.38	0.004	.0203866	.0453614
x13	-.0014987	.0016936	-0.88	0.441	-.0068886	.0038912
x14	.002428	.0035094	0.69	0.539	-.0087404	.0135963
x15	-.9407965	.8855088	-1.06	0.366	-3.758881	1.877288
x16	.0045435	.0041557	1.09	0.354	-.0086817	.0177688
x17	-.00346	.0027495	-1.26	0.297	-.01221	.0052901
x18	.0020247	.0033324	0.61	0.586	-.0085804	.0126299
x19	.01245	.0556645	0.22	0.837	-.1646994	.1895994
x20	-.002517	.000913	-2.76	0.070	-.0054225	.0003885
x21	-.0022353	.0008965	-2.49	0.088	-.0050885	.0006179
x22	.0051354	.1253395	0.04	0.970	-.3937509	.4040216
x23	-.000914	.0194429	-0.05	0.965	-.0627901	.060962
x24	-.0107727	.0089902	-1.20	0.317	-.0393834	.0178381
x25	.0019414	.0014501	1.34	0.273	-.0026733	.0065561
x26	.0008729	.000841	1.04	0.376	-.0018034	.0035493
x27	-.0010262	.0020618	-0.50	0.653	-.0075876	.0055352
x28	.0194033	.0158757	1.22	0.309	-.0311204	.069927
x29	-.0434677	.1110943	-0.39	0.722	-.3970193	.3100839
x30	-.0049231	.0010514	-4.68	0.018	-.008269	-.0015772
x31	.003375	.0027439	1.23	0.306	-.0053574	.0121075
x32	.0006721	.0010984	0.61	0.584	-.0028236	.0041678
x33	.0006039	.0038777	0.16	0.886	-.0117367	.0129444
x34	-.0003681	.0012216	-0.30	0.783	-.0042558	.0035196
x35	(omitted)					
_cons	1.366668	1.282227	1.07	0.365	-2.713952	5.447288
sigma_u	.10968697					
sigma_e	.267204					
rho	.14420887	(fraction of variance due to u_i)				

El modelo final obtenido es el siguiente:

```
. xtreg logproporción L.logproporción L4.logproporción x2 x7 x9 x10 x12 x30, fe v
> ce(cluster panel)
```

Fixed-effects (within) regression  
Group variable: panel

Number of obs = 416  
Number of groups = 4

R-sq: within = 0.2565  
between = 0.7913  
overall = 0.4541

Obs per group: min = 104  
avg = 104.0  
max = 104

corr(u\_i, Xb) = -0.0251

F(3,3) =  
Prob > F =

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
logproporc~n						
L1.	.3528319	.0191519	18.42	0.000	.2918819	.4137818
L4.	.1667528	.0565935	2.95	0.060	-.0133531	.3468587
x2	-1.539513	.3380018	-4.55	0.020	-2.615185	-.4638403
x7	-.0049989	.0009227	-5.42	0.012	-.0079354	-.0020624
x9	.0332369	.00143	23.24	0.000	.0286861	.0377876
x10	.0113771	.0010926	10.41	0.002	.0079001	.0148541
x12	.0366787	.0033515	10.94	0.002	.0260129	.0473446
x30	-.0063389	.0016969	-3.74	0.033	-.0117392	-.0009387
_cons	.0943342	.1730459	0.55	0.624	-.4563751	.6450434
sigma_u	.12334146					
sigma_e	.26715236					
rho	.17570456	(fraction of variance due to u_i)				

En este modelo las variables son todas ellas significativas al 10 % de nivel de significación, excepto la constante, pero que debemos retener.

Las variables incluidas son:

- 1. retardo1 y retardo4.** L1, L2. Recogen los valores de la variable explicativa desfasados uno y cuatro periodos. El signo positivo indica que estas variables y la variable a explicar evolucionan en el mismo sentido.
- 2. Diferencial de precios respecto al Principal Competidor.**  $x_2$ . En este caso el signo asociado es negativo, diferente del esperado que era positivo. Esto es, cuanto menor es la diferencia en precios con la marca competidora, mayores son las ventas de la Marca 1. Por lo tanto y como estrategia de precio en este formato, lo más lógico debería ser acercarse al precio de la competencia y no buscar situaciones donde la diferencia entre ellos sean muy notorias.

3. **Distribución Ponderada Marcas de la Distribución:**  $x_7$  Como tiene signo negativo, que es el esperado. Cuanto mayor es la distribución de las marcas de distribución menor es la cuota de mercado de la Marca 1.
4. **Cobertura Promoción Folleto y Exposición Especial Marca 1:**  $x_9$  El signo es positivo, esto es, valores elevados de la variable (gran cobertura de esta acción) se asocian a valores altos de la cuota en tienda.
5. **Cobertura Promoción sólo Exposición Especial Marca 1:**  $x_{10}$  Signo positivo, valores elevados de esta variable se corresponden con cuotas en tienda elevadas ( $y$ ). La utilización de esta acción de marketing supone un aumento en la cuota en tienda de la Marca 1.
6. **Cobertura Otras Promociones de la Marca 1:**  $x_{12}$  La puesta en marcha de Otras Promociones diferentes a las consideradas anteriormente, tiene de nuevo efectos positivos sobre la cuota en tienda de la Marca 1.
7. **Cobertura Promoción Folleto Y Exposición Especial Marcas Distribución:**  $x_{30}$  El signo asociado a esta variable es negativo, por lo tanto, coberturas elevadas en esta promoción por parte de las Marcas de Distribución, producen una reducción en la cuota de mercado de la marca.

El contraste conjunto nos permite comprobar que efectivamente las variables que hemos eliminado del modelo no son significativas.

```

. testparm retardo2logproporcionN retardo3logproporcionN x1 x3 x4 x5 x6 x8 x11 x13
> x14 x15 x16 x17 x18 x19 x20 x21 x22 x23 x24 x25 x26 x27 x28 x29 x31 x32 x33 x34
> x35

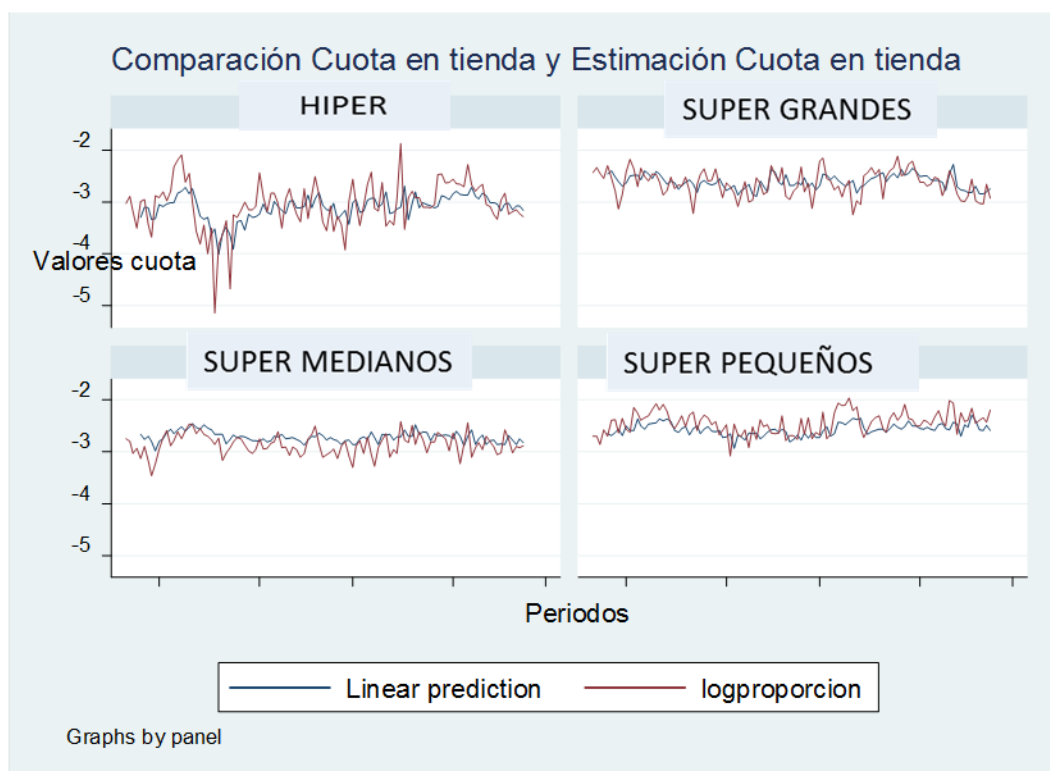
( 1) x1 = 0
( 2) x3 = 0
( 3) x4 = 0
( 4) x5 = 0
( 5) x6 = 0
( 6) x8 = 0
( 7) x11 = 0
( 8) x13 = 0
( 9) x14 = 0
(10) x15 = 0
(11) x16 = 0
(12) x17 = 0
(13) x18 = 0
(14) x19 = 0
(15) x20 = 0
(16) x21 = 0
(17) x22 = 0
(18) x23 = 0
(19) x24 = 0
(20) x25 = 0
(21) x26 = 0
(22) x27 = 0
(23) x28 = 0
(24) x29 = 0
(25) x31 = 0
(26) x32 = 0
(27) x33 = 0
(28) x34 = 0
      Constraint 1 dropped
      Constraint 3 dropped
      Constraint 4 dropped
      Constraint 5 dropped
      Constraint 7 dropped
      Constraint 8 dropped
      Constraint 9 dropped
      Constraint 11 dropped
      Constraint 12 dropped
      Constraint 13 dropped
      Constraint 14 dropped
      Constraint 15 dropped
      Constraint 16 dropped
      Constraint 17 dropped
      Constraint 18 dropped
      Constraint 19 dropped
      Constraint 20 dropped
      Constraint 21 dropped
      Constraint 22 dropped
      Constraint 23 dropped
      Constraint 24 dropped
      Constraint 25 dropped
      Constraint 26 dropped
      Constraint 27 dropped
      Constraint 28 dropped

      F( 3, 3) = 1.09
      Prob > F = 0.4728

```

El p valor asociado a esta prueba nos indica que debemos aceptar la hipótesis nula, esto es las variables aquí incluidas resultan ser no significativas y podemos eliminarlas del modelo.





Este gráfico y la lectura de los valores del R cuadrado que nos proporciona la salida del modelo nos hacen plantearnos nuevamente si el peso que tienen las distintas variables en el modelo es igual o depende del tipo de establecimiento como ya se ha visto que sucede en el formato anterior.

Para calcular el efecto de cada uno de los tipos de establecimientos se plantea una regresión de variable ficticia sobre la variable a explicar, las variables originales, tres variables ficticias (hipermercados, supermercados grandes y supermercados medianos; no se incluye supermercados pequeños porque es la categoría de referencia) y las interacciones de estas variables con las variables que se incluyeron en el modelo.

Los resultados son los que se muestran a continuación, algunas variables se han eliminado como consecuencia de la colinealidad que presentan con otras.

```
. regress logproporcion retardo1logproporcion retardo4logproporcionÑ x2 x7 x9 x10
> x12 x30 r1h r1sg r1sm r4h r4sg r4sm intx2hiper intx2superg intx2superm intx7hipe
> r intx7superg intx7superm intx9hiper intx9superg intx9superm intx10hiper intx10s
> uperg intx10superm intx12hiper intx12superg intx12superm intx30hiper intx30super
> g intx30superm hiper superg superm, vce(cluster panel)
note: intx9hiper omitted because of collinearity
note: intx9superg omitted because of collinearity
note: intx9superm omitted because of collinearity
note: intx10hiper omitted because of collinearity
note: intx10superm omitted because of collinearity
note: intx12hiper omitted because of collinearity
note: intx12superg omitted because of collinearity
note: intx12superm omitted because of collinearity
note: intx30superm omitted because of collinearity
```

Linear regression

```
Number of obs =      416
F( 0,      3) =      .
Prob > F       =
R-squared      = 0.5415
Root MSE      = 0.26925
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.2262146	2.51e-12	9.0e+10	0.000	.2262146	.2262146
retardo4lo~Ñ	.0673898	1.31e-13	5.1e+11	0.000	.0673898	.0673898
x2	-2.758414	5.24e-11	-5.3e+10	0.000	-2.758414	-2.758414
x7	-.0078564	5.12e-14	-1.5e+11	0.000	-.0078564	-.0078564
x9	.0326381	2.57e-13	1.3e+11	0.000	.0326381	.0326381
x10	.0094385	2.51e-13	3.8e+10	0.000	.0094385	.0094385
x12	.032236	1.74e-13	1.8e+11	0.000	.032236	.032236
x30	-.0061019	2.56e-13	-2.4e+10	0.000	-.0061019	-.0061019
r1h	.1358611	2.52e-12	5.4e+10	0.000	.1358611	.1358611
r1sg	.118955	2.93e-12	4.1e+10	0.000	.118955	.118955
r1sm	.0815062	2.59e-12	3.2e+10	0.000	.0815062	.0815062
r4h	.1718397	6.34e-13	2.7e+11	0.000	.1718397	.1718397
r4sg	-.0203512	1.48e-12	-1.4e+10	0.000	-.0203512	-.0203512
r4sm	-.0117483	2.96e-13	-4.0e+10	0.000	-.0117483	-.0117483
intx2hiper	1.542937	5.27e-11	2.9e+10	0.000	1.542937	1.542937
intx2superg	2.072904	8.06e-11	2.6e+10	0.000	2.072904	2.072904
intx2superm	.7730699	5.26e-11	1.5e+10	0.000	.7730699	.7730699
intx7hiper	-.000432	5.14e-14	-8.4e+09	0.000	-.000432	-.000432
intx7superg	.0022115	8.09e-14	2.7e+10	0.000	.0022115	.0022115
intx7superm	.0041327	5.16e-14	8.0e+10	0.000	.0041327	.0041327
intx9hiper	(omitted)					
intx9superg	(omitted)					
intx9superm	(omitted)					
intx10hiper	(omitted)					
intx10superg	.0000899	2.77e-13	3.2e+08	0.000	.0000899	.0000899
intx10superm	(omitted)					
intx12hiper	(omitted)					
intx12superg	(omitted)					
intx12superm	(omitted)					
intx30hiper	-.0281928	2.83e-13	-1.0e+11	0.000	-.0281928	-.0281928
intx30superg	-.0003483	2.65e-13	-1.3e+09	0.000	-.0003483	-.0003483
intx30superm	(omitted)					
hiper	-.4349196	4.92e-11	-8.8e+09	0.000	-.4349196	-.4349196
superg	-1.696912	6.96e-11	-2.4e+10	0.000	-1.696912	-1.696912
superm	-.9373742	4.88e-11	-1.9e+10	0.000	-.9373742	-.9373742
_cons	.7603466	4.88e-11	1.6e+10	0.000	.7603466	.7603466

Eliminado las variables que presentan colinealidad, nos queda el modelo definitivo.

```
. regress logproporción retardo1logproporción retardo4logproporción x2 x7 x9 x10
> x12 x30 r1h r1sg r1sm r4h r4sg r4sm intx2hiper intx2superg intx2superm intx7hipe
> r intx7superg intx7superm intx10superg intx30hiper intx30superg hiper superg sup
> erm, vce(cluster panel)
```

Linear regression

Number of obs = 416  
F( 0, 3) = .  
Prob > F = .  
R-squared = 0.5415  
Root MSE = .26925

(Std. Err. adjusted for 4 clusters in panel)

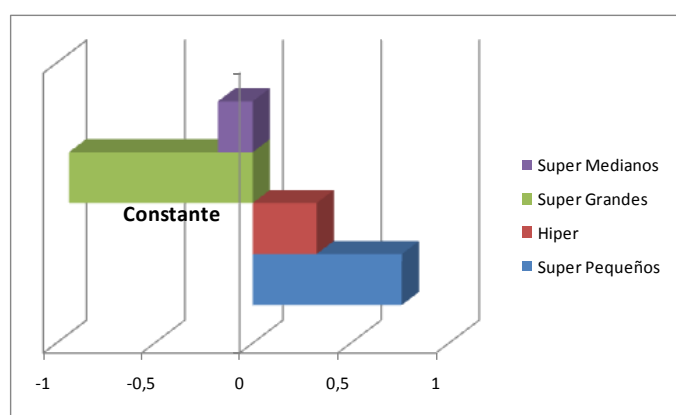
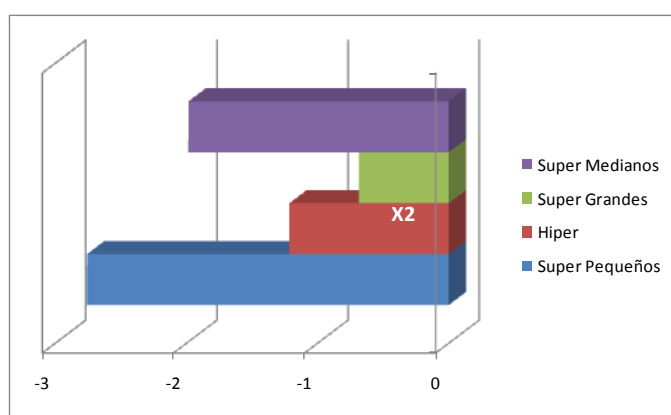
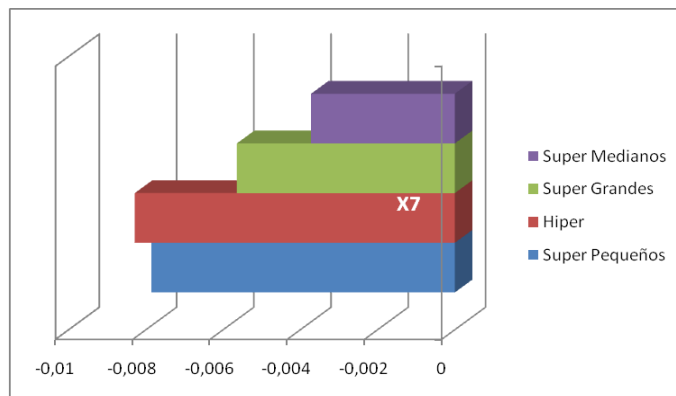
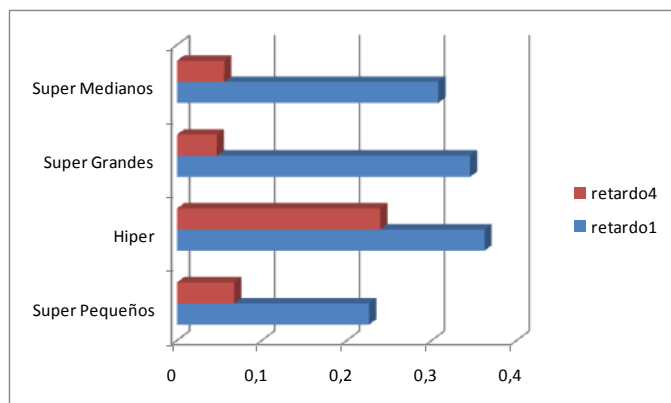
logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.2262146	2.51e-12	9.0e+10	0.000	.2262146	.2262146
retardo4lo~n	.0673898	1.31e-13	5.1e+11	0.000	.0673898	.0673898
x2	-2.758414	5.24e-11	-5.3e+10	0.000	-2.758414	-2.758414
x7	-.0078564	5.12e-14	-1.5e+11	0.000	-.0078564	-.0078564
x9	.0326381	2.57e-13	1.3e+11	0.000	.0326381	.0326381
x10	.0094385	2.51e-13	3.8e+10	0.000	.0094385	.0094385
x12	.032236	1.74e-13	1.8e+11	0.000	.032236	.032236
x30	-.0061019	2.56e-13	-2.4e+10	0.000	-.0061019	-.0061019
r1h	.1358611	2.52e-12	5.4e+10	0.000	.1358611	.1358611
r1sg	.118955	2.93e-12	4.1e+10	0.000	.118955	.118955
r1sm	.0815062	2.59e-12	3.2e+10	0.000	.0815062	.0815062
r4h	.1718397	6.34e-13	2.7e+11	0.000	.1718397	.1718397
r4sg	-.0203512	1.48e-12	-1.4e+10	0.000	-.0203512	-.0203512
r4sm	-.0117483	2.96e-13	-4.0e+10	0.000	-.0117483	-.0117483
intx2hiper	1.542937	5.27e-11	2.9e+10	0.000	1.542937	1.542937
intx2superg	2.072904	8.06e-11	2.6e+10	0.000	2.072904	2.072904
intx2superm	.7730699	5.26e-11	1.5e+10	0.000	.7730699	.7730699
intx7hiper	-.000432	5.14e-14	-8.4e+09	0.000	-.000432	-.000432
intx7superg	.0022115	8.09e-14	2.7e+10	0.000	.0022115	.0022115
intx7superm	.0041327	5.16e-14	8.0e+10	0.000	.0041327	.0041327
intx10superg	.0000899	2.77e-13	3.2e+08	0.000	.0000899	.0000899
intx30hiper	-.0281928	2.83e-13	-1.0e+11	0.000	-.0281928	-.0281928
intx30superg	-.0003483	2.65e-13	-1.3e+09	0.000	-.0003483	-.0003483
hiper	-.4349196	4.92e-11	-8.8e+09	0.000	-.4349196	-.4349196
superg	-1.696912	6.96e-11	-2.4e+10	0.000	-1.696912	-1.696912
superm	-.9373742	4.88e-11	-1.9e+10	0.000	-.9373742	-.9373742
_cons	.7603466	4.88e-11	1.6e+10	0.000	.7603466	.7603466

Los efectos de las variables incluidas en el modelo según los tipos de establecimientos son:

	Super Pequeños	Hiper	Super Grandes	Super Medianos
retardo1	0,2262146	0,3620757	0,3451696	0,3077208
retardo4	0,0673898	0,2392295	0,0470386	0,0556415
X <sub>2</sub>	-2,758414	-1,215477	-0,68551	-1,9853441
X <sub>7</sub>	-0,0078564	-0,0082884	-0,0056449	-0,0037237
X <sub>9</sub>	0,0326381			
X <sub>10</sub>	0,0094385		0,0095284	
X <sub>12</sub>	0,032236			
X <sub>30</sub>	-0,0061019	-0,0342947	-0,0064502	
Constante	0,7603466	0,325427	-0,9365654	-0,1770276

En las variables x<sub>9</sub>, x<sub>10</sub>, x<sub>12</sub> y x<sub>30</sub> no se puede calcular el efecto en los distintos tipos de establecimientos por los problemas de colinealidad que presenta el modelo, así en las variables x<sub>9</sub>, y x<sub>12</sub> resulta imposible calcular el efecto para cualquier tipo de

establecimiento, mientras que en las variables  $x_{10}$  y  $x_{30}$  sólo pueden calcularse los efectos de algún tipo de establecimiento.



Para cerrar este apartado vamos a realizar el contraste Reset para determinar si la forma funcional es la adecuada.

Hacemos una regresión sobre la variable original y utilizamos como variables explicativas las variables significativas en los modelos anteriores más la variable estimada al cuadrado (pred2 es el nombre de la nueva variable incluida), a partir de este momento se trata de ver si la nueva variable es significativa o no; por lo que realizamos un contraste tipo t sobre dicha variable.

Como puede apreciarse en la salida del programa, la variable pred2 resulta ser significativa a un 10 % de nivel de significación. Esta conclusión se ve corroborada

con el contraste tipo F que viene unas líneas más adelante. En este contraste puede apreciarse que la hipótesis nula **se rechaza** por lo que concluimos que esta variable aporta algo significativo al modelo y nos dice que la forma funcional del modelo no es correcta, ya que falta algo de orden 2 que no está incorporado al modelo.

```
. xtreg logproporción retardo1logproporción retardo4logproporción x2 x7 x9 x
> 10 x12 x30 pred2, fe vce(cluster panel )
```

Fixed-effects (within) regression  
Group variable: panel

Number of obs = 416  
Number of groups = 4

R-sq: within = 0.2575  
between = 0.7713  
overall = 0.4473

Obs per group: min = 104  
avg = 104.0  
max = 104

corr(u\_i, xb) = -0.0308

F(3,3) = .  
Prob > F = .

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.1503918	.0685394	2.19	0.116	-.0677312	.3685148
retardo4lo~n	.0694991	.0856475	0.81	0.477	-.2030695	.3420676
x2	-.6672647	.3348917	-1.99	0.140	-1.73304	.3985101
x7	-.0023706	.001472	-1.61	0.206	-.0070553	.002314
x9	.0168191	.0048733	3.45	0.041	.0013102	.032328
x10	.0045809	.0025625	1.79	0.172	-.0035741	.0127359
x12	.019357	.0091018	2.13	0.123	-.0096089	.0483229
x30	-.0033805	.0027729	-1.22	0.310	-.0122051	.0054441
pred2	-.0954206	.0338788	-2.82	0.067	-.2032379	.0123968
_cons	-.7959385	.335227	-2.37	0.098	-1.86278	.2709033
sigma_u	.12915567					
sigma_e	.26730524					
rho	.18927229	(fraction of variance due to u_i)				

```
. testparm pred2
```

( 1) pred2 = 0

F( 1, 3) = 7.93  
Prob > F = 0.0669

Por lo tanto, hay que buscar que no linealidades están presentes y que no han sido incluidas en la ecuación del modelo.

Los resultados de esta búsqueda se muestran en la salida siguiente; en ella puede apreciarse que se eliminan algunas variables, siendo su efecto recogido por la interacción con otras, otras se incorporan con exponentes al cuadrado y otras con exponentes al cubo.

```
. xtreg logproporcion retardo1logproporcion retardo4logproporcionN x2 x7 x10
> x30 x9cuadrado x7x9 x7x12 retardo4cuadrado x2cubo x9cubo , fe vce(cluster
> panel )
```

Fixed-effects (within) regression  
 Group variable: panel

Number of obs = 416  
 Number of groups = 4

R-sq: within = 0.2864  
 between = 0.6931  
 overall = 0.4356

obs per group: min = 104  
 avg = 104.0  
 max = 104

corr(u\_i, Xb) = -0.1209

F(3,3) = .  
 Prob > F = .

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.3580636	.020344	17.60	0.000	.29332	.4228073
retardo4lo~N	-.817763	.1170411	-6.99	0.006	-1.19024	-.445286
x2	-8.563382	2.488711	-3.44	0.041	-16.48357	-.6431919
x7	-.0056555	.0011112	-5.09	0.015	-.0091918	-.0021193
x10	.0082547	.0005385	15.33	0.001	.0065411	.0099683
x30	-.0050432	.0007741	-6.52	0.007	-.0075066	-.0025798
x9cuadrado	-.0724968	.0020161	-35.96	0.000	-.0789128	-.0660808
x7x9	.0065381	.0002281	28.66	0.000	.0058122	.007264
x7x12	.0005653	.0000425	13.31	0.001	.0004301	.0007004
retardo4cu~o	-.1627271	.0173756	-9.37	0.003	-.218024	-.1074301
x2cubo	3.506264	1.158777	3.03	0.056	-.1814811	7.19401
x9cubo	.0054587	.0001424	38.33	0.000	.0050054	.005912
_cons	2.492593	1.414016	1.76	0.176	-2.007437	6.992624
sigma_u	.15127578					
sigma_e	.26304201					
rho	.24853922	(fraction of variance due to u_i)				

Para finalizar aplicamos sobre esta especificación el contraste Reset obteniendo los resultados siguientes:

```
. xtreg logproporcion retardo1logproporcion retardo4logproporcionN x2 x7 x10
> x30 x9cuadrado x7x9 x7x12 retardo4cuadrado x2cubo x9cubo pred52 , fe vce(
> cluster panel )
```

Fixed-effects (within) regression  
 Group variable: panel

Number of obs = 416  
 Number of groups = 4

R-sq: within = 0.2865  
 between = 0.6987  
 overall = 0.4379

obs per group: min = 104  
 avg = 104.0  
 max = 104

corr(u\_i, Xb) = -0.1160

F(3,3) = .  
 Prob > F = .

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.4382561	.1788927	2.45	0.092	-.1310603	1.007573
retardo4lo~N	-1.053066	.5693231	-1.85	0.161	-2.864906	.7587742
x2	-10.01269	4.273265	-2.34	0.101	-23.61213	3.586741
x7	-.0067775	.0028604	-2.37	0.099	-.0158805	.0023256
x10	.009928	.0028875	3.44	0.041	.0007388	.0191172
x30	-.0059804	.0024923	-2.40	0.096	-.0139119	.0019511
x9cuadrado	-.0870219	.0298814	-2.91	0.062	-.1821177	.008074
x7x9	.0078419	.0027093	2.89	0.063	-.0007805	.0164642
x7x12	.0006656	.0002007	3.32	0.045	.000027	.0013042
retardo4cu~o	-.2086942	.1063295	-1.96	0.144	-.5470822	.1296938
x2cubo	4.067305	1.863072	2.18	0.117	-1.861823	9.996433
x9cubo	.0065501	.0022326	2.93	0.061	-.0005551	.0136553
pred52	.0373752	.0729405	0.51	0.644	-.194754	.2695044
_cons	3.046782	1.940341	1.57	0.214	-3.12825	9.221815
sigma_u	.14973449					
sigma_e	.26335095					
rho	.24429994	(fraction of variance due to u_i)				

```
. testparm pred52
```

```
( 1) pred52 = 0
```

F( 1, 3) = 0.26  
 Prob > F = 0.6437

En el contraste se acepta la hipótesis nula; parece por tanto que hasta aquí la forma funcional que une las variables es correcta.

```
. xtreg logproporcion retardo1logproporcion retardo4logproporcion x2 x7 x10
> x30 x9cuadrado x7x9 x7x12 retardo4cuadrado x2cubo x9cubo pred52 pred53,
> fe vce(cluster panel)
```

```
Fixed-effects (within) regression              Number of obs   =       416
Group variable: panel                        Number of groups =         4

R-sq:  within = 0.2866                      Obs per group: min =       104
        between = 0.6918                      avg =      104.0
        overall = 0.4350                      max =       104
```

```
corr(u_i, Xb) = -0.1390                      F(3,3) = .
                                                Prob > F = .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	-.0761748	1.313233	-0.06	0.957	-4.25547	4.10312
retardo4lo~n	.0892009	2.758838	0.03	0.976	-8.690654	8.869055
x2	1.946675	30.03493	0.06	0.952	-93.63789	97.53124
x7	.0012573	.0207097	0.06	0.955	-.0646501	.0671647
x10	-.0019935	.0313322	-0.06	0.953	-.1017067	.0977196
x30	.0011727	.0183777	0.06	0.953	-.0573133	.0596588
x9cuadrado	.0151784	.2634193	0.06	0.958	-.8231394	.8534962
x7x9	-.0013831	.0237823	-0.06	0.957	-.0770688	.0743027
x7x12	-.0001256	.0020974	-0.06	0.956	-.0068004	.0065491
retardo4cu~o	.019316	.5573406	0.03	0.975	-1.754391	1.793022
x2cubo	-.8070111	12.17623	-0.07	0.951	-39.55721	37.94319
x9cubo	-.00114	.0198276	-0.06	0.958	-.0642404	.0619603
pred52	-.4312933	1.253524	-0.34	0.754	-4.420566	3.55798
pred53	-.0501553	.1393914	-0.36	0.743	-.4937609	.3934503
_cons	-1.83276	12.28662	-0.15	0.891	-40.93426	37.26874
sigma_u	.15215384					
sigma_e	.26366377					
rho	.24982137	(fraction of variance due to u_i)				

```
. testparm pred52 pred53
```

```
( 1) pred52 = 0
( 2) pred53 = 0
```

```
F( 2, 3) = 0.17
Prob > F = 0.8522
```

Hipótesis que se ve confirmada con este contraste, y por lo tanto podemos decir que esta forma funcional es la adecuada para estas variables del modelo.

Por tanto el modelo estimado definitivo queda como sigue:

```
. xtreg logproporcion retardo1logproporcion retardo4logproporcion x2 x7 x10 x30 x9cuadra
> do x7x9 x7x12 retardo4cuadrado x2cubo x9cubo, fe vce(cluster panel)
```

```
Fixed-effects (within) regression               Number of obs   =       416
Group variable: panel                          Number of groups =        4

R-sq:  within = 0.2864                        Obs per group:  min =       104
        between = 0.6931                      avg           =    104.0
        overall = 0.4356                      max           =       104
```

```
corr(u_i, Xb)  = -0.1209                      F(3,3)          =        .
                                                Prob > F        =        .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc-n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo-n	.3580636	.020344	17.60	0.000	.29332	.4228073
retardo4lo-n	-.817763	.1170411	-6.99	0.006	-1.19024	-.445286
x2	-8.563382	2.488711	-3.44	0.001	-16.48357	-.6431919
x7	-.0056555	.0011112	-5.09	0.015	-.0091918	-.0021193
x10	.0082547	.0005385	15.33	0.001	.0065411	.0099683
x30	-.0050432	.0007741	-6.52	0.007	-.0075066	-.0025798
x9cuadrado	-.0724968	.0020161	-35.96	0.000	-.0789128	-.0660808
x7x9	.0065381	.0002281	28.66	0.000	.0058122	.007264
x7x12	.0005653	.0000425	13.31	0.001	.0004301	.0007004
retardo4cu-o	-.1627271	.0173756	-9.37	0.003	-.218024	-.1074301
x2cubo	3.506264	1.158777	3.03	0.056	-.1814811	7.19401
x9cubo	.0054587	.0001424	38.33	0.000	.0050054	.005912
_cons	2.492593	1.414016	1.76	0.176	-2.007437	6.992624
sigma_u	.15127578					
sigma_e	.26304201					
rho	.24853922	(fraction of variance due to u_i)				



### 6.3.-ESTIMACIÓN DEL MODELO DINÁMICO DE EFECTOS INDIVIDUALES ESPECÍFICOS FORMATO DE 500 A 899 g.

La estimación del modelo de partida es la que sigue:

```
. xtreg logproporcion retardo1logproporcion retardo2logproporcion retardo3logprop
> orcion retardo4logproporcion x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 x13 x14 x1
> 5 x16 x17 x18 x19 x20 x21 x22 x23 x24 x25 x26 x27 x28 x29 x30 x31 x32 x33 x34 x3
> 5, fe vce(cluster panel)
note: x33 omitted because of collinearity
note: x35 omitted because of collinearity
```

Fixed-effects (within) regression  
Group variable: panel

Number of obs = 416  
Number of groups = 4

R-sq: within = 0.5574  
between = 0.9920  
overall = 0.9508

Obs per group: min = 104  
avg = 104.0  
max = 104

corr(u\_i, Xb) = 0.9364

F(3,3) = .  
Prob > F = .

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.4127647	.0969089	4.26	0.024	.1043572	.7211722
retardo2lo~Ñ	-.0994329	.0083661	-11.89	0.001	-.1260577	-.0728082
retardo3lo~Ñ	.0705416	.0489	1.44	0.245	-.08508	.2261631
retardo4lo~Ñ	.0524646	.0260155	2.02	0.137	-.0303282	.1352575
x1	.0020235	.0013358	1.51	0.227	-.0022278	.0062747
x2	.0609496	.1812074	0.34	0.759	-.5157332	.6376324
x3	.1659439	.1260041	1.32	0.279	-.2350573	.5669452
x4	.8684815	.1646613	5.27	0.013	.3444558	1.392507
x5	-.0016503	.0016551	-1.00	0.392	-.0069177	.0036172
x6	-.0015904	.0023536	-0.68	0.548	-.0090805	.0058997
x7	.0001935	.0010555	0.18	0.866	-.0031657	.0035526
x8	.30535	.3296821	0.93	0.423	-.7438456	1.354546
x9	.0061512	.0011013	5.59	0.011	.0026465	.0096559
x10	.0059483	.0039326	1.51	0.228	-.0065671	.0184636
x11	.0009593	.0010358	0.93	0.423	-.002337	.0042555
x12	-.0014447	.0002962	-4.88	0.016	-.0023873	-.000502
x13	.0011179	.0012344	0.91	0.432	-.0028104	.0050462
x14	.000907	.0078982	0.11	0.916	-.0242285	.0260425
x15	.6742674	.2224114	3.03	0.056	-.0335449	1.38208
x16	.0014933	.0021373	0.70	0.535	-.0053085	.008295
x17	.0006291	.0018541	0.34	0.757	-.0052713	.0065295
x18	.0003042	.0003595	0.85	0.460	-.0008398	.0014483
x19	-.0003695	.001842	-0.20	0.854	-.0062317	.0054927
x20	-.0001671	.000579	-0.29	0.792	-.0020097	.0016755
x21	.0006574	.000684	0.96	0.407	-.0015194	.0028342
x22	-.0321163	.1114651	-0.29	0.792	-.3868481	.3226154
x23	-.0025252	.0026908	-0.94	0.417	-.0110886	.0060381
x24	-.0015699	.0001462	-10.74	0.002	-.0020353	-.0011046
x25	.000134	.0009295	0.14	0.894	-.0028242	.0030923
x26	.0011023	.0011793	0.93	0.419	-.0026508	.0048554
x27	.0001805	.0004274	0.42	0.701	-.0011796	.0015407
x28	.0051723	.0024895	2.08	0.129	-.0027505	.013095
x29	-.1511497	.02818	-5.36	0.013	-.2408309	-.0614684
x30	.0010308	.0030576	0.34	0.758	-.0086998	.0107613
x31	-.001158	.0002551	-4.54	0.020	-.0019697	-.0003462
x32	.0017978	.0018566	0.97	0.404	-.0041108	.0077065
x33	(omitted)					
x34	-.0004162	.0013678	-0.30	0.781	-.0047692	.0039369
x35	(omitted)					
_cons	-2.390478	.6360198	-3.76	0.033	-4.414577	-.3663791
sigma_u	.3379415					
sigma_e	.09073843					
rho	.93275399	(fraction of variance due to u_i)				

Se observa que existen una gran cantidad de variables no significativas que es preciso eliminar para obtener un modelo más sencillo.

El modelo final, una vez eliminadas las variables no significativas es el siguiente:

```
. xtreg logproporcion retardollogproporcion x4 x9 x10 x15, fe vce(cluster panel)
```

Fixed-effects (within) regression  
Group variable: panel

Number of obs = 428  
Number of groups = 4

R-sq: within = 0.5440  
between = 0.9989  
overall = 0.9520

Obs per group: min = 107  
avg = 107.0  
max = 107

corr(u\_i, xb) = 0.9451  $F(3,3)$  = .  
Prob > F = .

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardollogproporc~n	.483284	.024683	19.58	0.000	.4047317	.5618362
x4	.803612	.3188837	2.52	0.086	-.2112182	1.818442
x9	.0089608	.0011252	7.96	0.004	.00538	.0125416
x10	.0075775	.0030303	2.50	0.088	-.0020662	.0172212
x15	.3612625	.201384	1.79	0.171	-.2796314	1.002156
_cons	-1.742882	.1687559	-10.33	0.002	-2.279938	-1.205825
sigma_u	.33765923					
sigma_e	.09484762					
rho	.92686706	(fraction of variance due to u_i)				

En este modelo todas las variables incluidas son significativas al menos al 10 %, excepto x15, la razón por la que esta variable se ha mantenido es que aunque individualmente es no significativa, cuando realizamos el contraste conjunto, que veremos a continuación, rechazábamos la hipótesis nula de No significatividad; posiblemente debido a un problema de multicolinealidad. Por tanto es preferible incluirla en el modelo y al realizar el contraste conjunto No RECHAZAMOS la hipótesis nula de NO SIGNIFICATIVIDAD, obteniendo un modelo más sencillo.

```

. testparm retardo2logproporcionN retardo3logproporcionN retardo4logproporcionN x1
> x2 x3 x5 x6 x7 x8 x11 x12 x13 x14 x16 x17 x18 x19 x20 x21 x22 x23 x24 x25 x26 x
> 27 x28 x29 x30 x31 x32 x33 x34 x35

( 1) retardo2logproporcionN = 0
( 2) retardo3logproporcionN = 0
( 3) retardo4logproporcionN = 0
( 4) x1 = 0
( 5) x2 = 0
( 6) x3 = 0
( 7) x5 = 0
( 8) x6 = 0
( 9) x7 = 0
(10) x8 = 0
(11) x11 = 0
(12) x12 = 0
(13) x13 = 0
(14) x14 = 0
(15) x16 = 0
(16) x17 = 0
(17) x18 = 0
(18) x19 = 0
(19) x20 = 0
(20) x21 = 0
(21) x22 = 0
(22) x23 = 0
(23) x24 = 0
(24) x25 = 0
(25) x26 = 0
(26) x27 = 0
(27) x28 = 0
(28) x29 = 0
(29) x30 = 0
(30) x31 = 0
(31) x32 = 0
(32) x34 = 0
Constraint 1 dropped
Constraint 2 dropped
Constraint 3 dropped
Constraint 4 dropped
Constraint 7 dropped
Constraint 8 dropped
Constraint 9 dropped
Constraint 11 dropped
Constraint 12 dropped
Constraint 13 dropped
Constraint 14 dropped
Constraint 15 dropped
Constraint 16 dropped
Constraint 17 dropped
Constraint 18 dropped
Constraint 19 dropped
Constraint 20 dropped
Constraint 21 dropped
Constraint 22 dropped
Constraint 23 dropped
Constraint 24 dropped
Constraint 25 dropped
Constraint 26 dropped
Constraint 27 dropped
Constraint 28 dropped
Constraint 29 dropped
Constraint 30 dropped
Constraint 31 dropped
Constraint 32 dropped

F( 3, 3) = 1.17
Prob > F = 0.4495

```

Por tanto, las variables que permanecen en el modelo son:

1. **Retardo1** Esta variable mide la influencia que tiene la variable a explicar (cuota de mercado) del periodo anterior. El signo asociado es positivo de tal forma que un aumento de la cuota de mercado en el periodo anterior produce un aumento en la variable a explicar del periodo actual (cuota de mercado).

**2. Diferencial Precio Marcas Distribución:**  $x_4$  A mayor diferencial de precio, mayor será la cuota en tienda de la Marca 1. El parámetro  $\beta$  asociado, cumple con la hipótesis de partida, es positivo; de forma que, incrementos en la variable  $x_4$  suponen un aumento del mismo signo en la variable explicativa.

**3. Cobertura promoción Folleto y Exposición Especial Marca 1:**  $x_9$  Esta variable mide la proporción de establecimientos que realizan esta acción conjunta. Su signo asociado es positivo, de manera que valores elevados de la variable (gran cantidad de establecimientos que realizan esta acción) se asocian a valores elevados de la cuota en tienda.

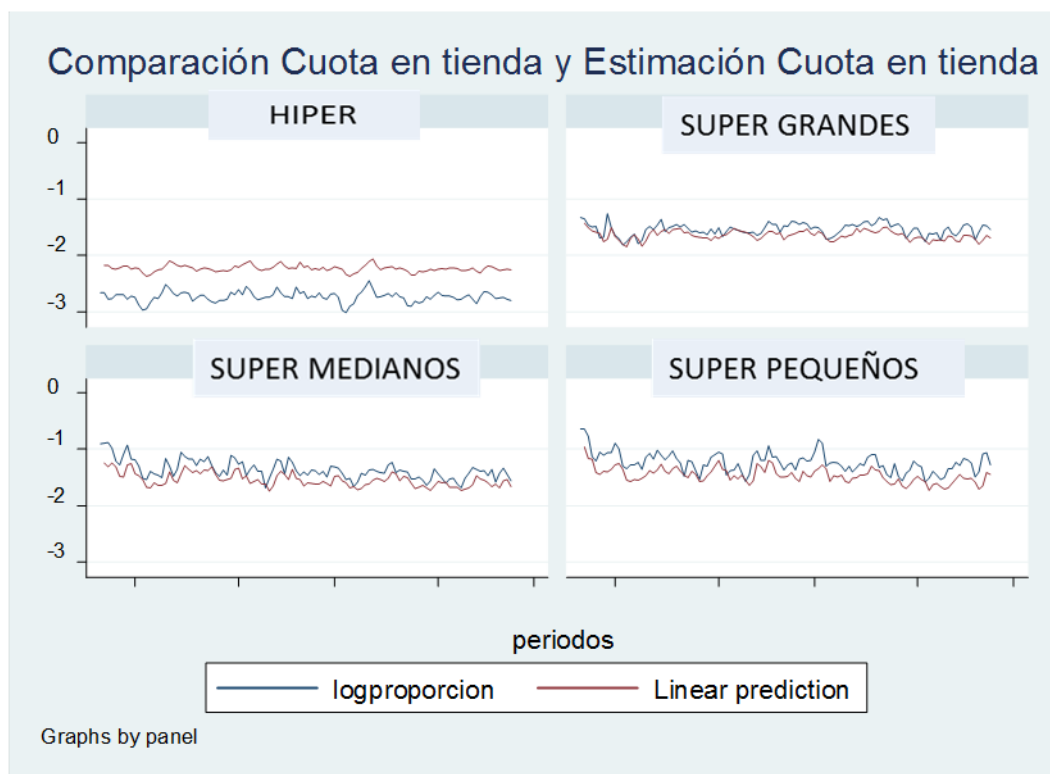
**4. Cobertura promoción SÓLO Exposición Especial Marca 1:**  $x_{10}$  La interpretación y el signo asociado a esta variable es el mismo que el de la variable  $x_9$ . Una cuestión que pone de manifiesto el modelo es que es la Exposición Especial la tiene un efecto dinamizador de la cuota en tienda, tanto cuando se utiliza individualmente o en conjunto con el folleto; ahora bien, el folleto sólo, no produce ningún tipo de reacción en la cuota de la marca 1.

**5. Intensidad de la Promoción de Precios del Principal Competidor:**  $x_{15}$  En el caso de esta variable el signo es positivo, cuando se espera un signo negativo. Ahora bien, si miramos el contraste de significación individual asociado a esta variable, resulta no ser significativa y debería eliminarse del modelo.

El siguiente paso es comprobar si el efecto de estas variables es el mismo para cada tipo de establecimiento, o como viene siendo habitual es distinto.

Una primera aproximación es el gráfico de la variable real y estimada por los distintos tipos de establecimientos.

En el gráfico que se muestra a continuación, puede observarse que en los Hiper existen unas mayores diferencias entre la variable real y la estimada, mientras que en los restantes paneles las diferencias no son tan evidentes; esta diferencia de comportamiento sugiere que el peso que tienen las variables explicativas no sea el mismo para todos los paneles (tipos de establecimientos).



Para comprobar si efectivamente los efectos de las variables en los establecimientos son distintos, procedemos como en los modelos anteriores; hacemos una regresión con las variables originales, tres variables ficticias (Hiper, Super Grandes y Super Medianos) y las interacciones entre estas variables ficticias y las variables originales. Los resultados se muestran a continuación.

En la salida del programa se ve que todas las variables incluidas son significativas, ya que los contraste  $t$  muestran un  $p$  valor cero; por lo tanto los efectos de las variables para cada tipo de establecimientos son diferentes.

```
. regress logproporcion retardo1logproporcion x4 x9 x10 x15 r1h r1superg r1superm
> inx4h inx4superg inx4superm inx9h inx9superg inx9superm inx10h inx10superg inx1
> 0superm inx15h inx15superg inx15superm hiper superg superm, vce(cluster panel)
```

Linear regression

Number of obs = 428  
F( 0, 3) = .  
Prob > F = .  
R-squared = 0.9779  
Root MSE = .09348

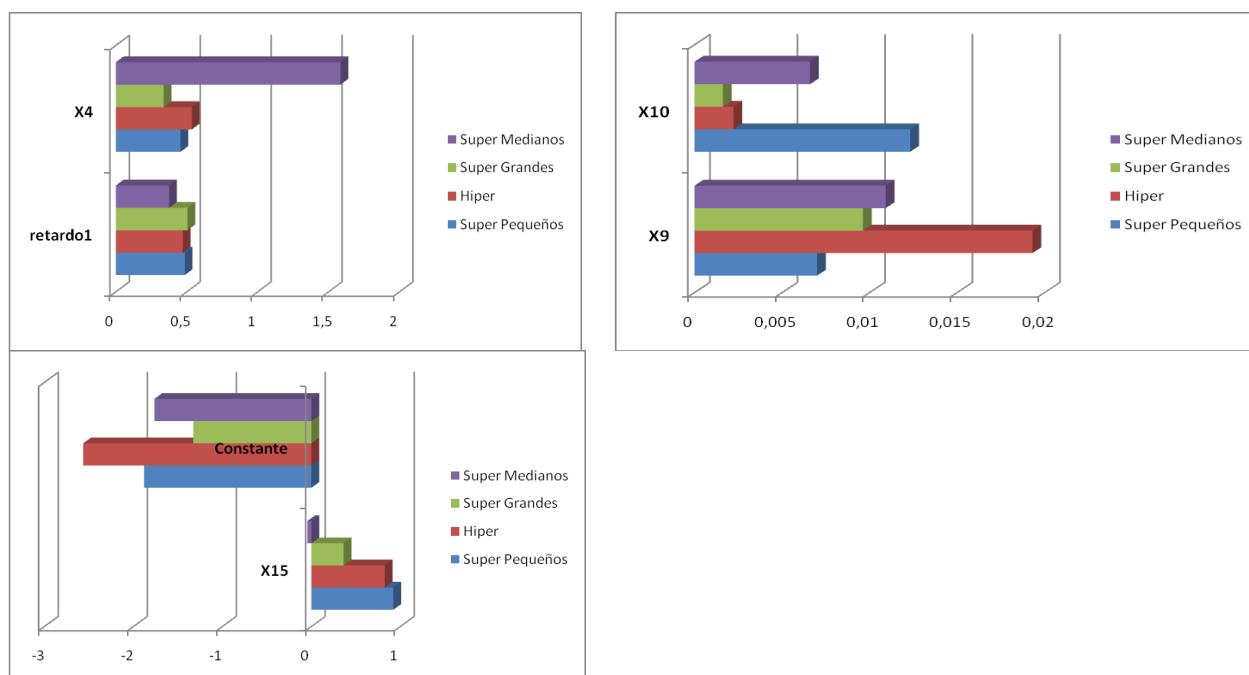
(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.4861747	1.48e-12	3.3e+11	0.000	.4861747	.4861747
x4	.4562307	2.66e-11	1.7e+10	0.000	.4562307	.4562307
x9	.0070006	8.09e-15	8.7e+11	0.000	.0070006	.0070006
x10	.0123224	4.10e-14	3.0e+11	0.000	.0123224	.0123224
x15	.919369	3.43e-11	2.7e+10	0.000	.919369	.919369
r1h	-.0139869	6.66e-12	-2.1e+09	0.000	-.0139869	-.0139869
r1superg	.0194782	1.60e-12	1.2e+10	0.000	.0194782	.0194782
r1superm	-.1103804	2.28e-12	-4.8e+10	0.000	-.1103804	-.1103804
inx4h	.0795798	6.79e-11	1.2e+09	0.000	.0795798	.0795798
inx4superg	-.1182929	2.69e-11	-4.4e+09	0.000	-.1182929	-.1182929
inx4superm	1.128388	3.01e-11	3.7e+10	0.000	1.128388	1.128388
inx9h	.0123084	4.46e-14	2.8e+11	0.000	.0123084	.0123084
inx9superg	.0026225	1.02e-14	2.6e+11	0.000	.0026225	.0026225
inx9superm	.0039224	1.47e-14	2.7e+11	0.000	.0039224	.0039224
inx10h	-.0100972	1.80e-13	-5.6e+10	0.000	-.0100972	-.0100972
inx10superg	-.0107093	4.19e-14	-2.6e+11	0.000	-.0107093	-.0107093
inx10superm	-.0057243	4.12e-14	-1.4e+11	0.000	-.0057243	-.0057243
inx15h	-.09571	3.91e-11	-2.4e+09	0.000	-.09571	-.09571
inx15superg	-.5582246	3.46e-11	-1.6e+10	0.000	-.5582246	-.5582246
inx15superm	-.9660698	3.59e-11	-2.7e+10	0.000	-.9660698	-.9660698
hiper	-.6853138	8.82e-11	-7.8e+09	0.000	-.6853138	-.6853138
superg	.5524935	5.07e-11	1.1e+10	0.000	.5524935	.5524935
superm	.1152333	5.42e-11	2.1e+09	0.000	.1152333	.1152333
_cons	-1.880785	5.02e-11	-3.7e+10	0.000	-1.880785	-1.880785

Los pesos de cada variable en los distintos establecimientos se recogen en esta tabla.

	Super Pequeños	Hiper	Super Grandes	Super Medianos
<i>retardo1</i>	0,4861747	0,4721878	0,5056529	0,3757943
<i>X<sub>4</sub></i>	0,4562307	0,5358105	0,3379378	1,5846187
<i>X<sub>9</sub></i>	0,0070006	0,019309	0,0096231	0,010923
<i>X<sub>10</sub></i>	0,0123224	0,0022252	0,0016131	0,0065981
<i>X<sub>15</sub></i>	0,919369	0,823659	0,3611444	-0,0467008
Constante	-1,880785	-2,5660988	-1,3282915	-1,7655517

En los gráficos se muestran los efectos de una forma más visual.



Por último queda realizar el contraste Reset para comprobar si la forma funcional utilizada es correcta; para ello procedemos como en los epígrafes anteriores, esto es, estimamos el modelo y obtenemos la variable estimada, para después calcular la variable estimada al cuadrado y al cubo. Posteriormente se introducen estas dos nuevas variables en el modelo y se calcula nuevamente la regresión; si las variables resultan ser no significativas, el contraste Reset es concluyente y la forma funcional es correcta.

Cuando se realiza la primera regresión del contraste Reset no aparece nada preocupante, la variable estimada al cuadrado no es significativa. Pero cuando damos el siguiente paso, esto es introducir como variable explicativa al cubo como variable explicativa, comienzan los problemas. Las variables estimadas al cuadrado y al cubo, introducidas como variables explicativas, resultan ser variables significativas; el test conjunto para comprobar si efectivamente ambas variables son cero nos dice que la hipótesis debe ser rechazada, por lo tanto el contraste Reset es concluyente, la forma funcional utilizada no es correcta. Esto es, existe alguna no linealidad en la forma funcional que no está recogida en el modelo.

```
. xtreg logproporcion retardo1logproporcion x4 x9 x10 x15 prediccionfe2, fe vce(c1
> uster panel )
```

```
Fixed-effects (within) regression      Number of obs   =      428
Group variable: panel                  Number of groups  =        4

R-sq:  within = 0.5440                  Obs per group:  min =      107
      between = 0.9988                      avg   =     107.0
      overall  = 0.9513                      max   =      107
```

```
corr(u_i, Xb)  = 0.9443                  F(3,3)          =      .
                                          Prob > F         =      .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.5000189	.1709639	2.92	0.061	-.0440644	1.044102
x4	.8307909	.3475503	2.39	0.097	-.2752693	1.936851
x9	.0092536	.0026304	3.52	0.039	.0008823	.0176248
x10	.0078126	.0041794	1.87	0.158	-.0054881	.0211132
x15	.3786073	.195264	1.94	0.148	-.2428099	1.000024
prediccion~2	.0109313	.1133214	0.10	0.929	-.3497081	.3715707
_cons	-1.781307	.2614192	-6.81	0.006	-2.613259	-.9493538
sigma_u	.33989571					
sigma_e	.09495976					
rho	.92759844	(fraction of variance due to u_i)				

```
. testparm prediccionfe2
```

```
( 1) prediccionfe2 = 0
```

```
F( 1, 3) = 0.01
Prob > F = 0.9292
```

```
. xtreg logproporcion retardo1logproporcion x4 x9 x10 x15 prediccionfe2 prediccion
> fe3, fe vce(cluster panel )
```

```
Fixed-effects (within) regression      Number of obs   =      428
Group variable: panel                  Number of groups  =        4

R-sq:  within = 0.5497                  Obs per group:  min =      107
      between = 0.9986                      avg   =     107.0
      overall  = 0.9453                      max   =      107
```

```
corr(u_i, Xb)  = 0.9376                  F(3,3)          =      .
                                          Prob > F         =      .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	2.37182	.3025447	7.84	0.004	1.408987	3.334652
x4	4.00527	.7177314	5.58	0.011	1.721129	6.289412
x9	.0440397	.0070308	6.26	0.008	.0216647	.0664148
x10	.0374871	.0023601	15.88	0.001	.0299761	.044998
x15	1.824947	.22825	8.00	0.004	1.098554	2.55134
prediccion~2	2.307393	.3364576	6.86	0.006	1.236634	3.378151
prediccion~3	.4374649	.0582064	7.52	0.005	.2522261	.6227038
_cons	-6.513809	.7425891	-8.77	0.003	-8.877058	-4.150559
sigma_u	.3596819					
sigma_e	.09447772					
rho	.93545751	(fraction of variance due to u_i)				

```
. testparm prediccionfe2 prediccionfe3
```

```
( 1) prediccionfe2 = 0
```

```
( 2) prediccionfe3 = 0
```

```
F( 2, 3) = 99.78
Prob > F = 0.0018
```



Por tanto el modelo estimado es el que se muestra en el cuadro siguiente y que es el primero que habíamos estimado:

Fixed-effects (within) regression	Number of obs	=	428
Group variable: panel	Number of groups	=	4
R-sq: within = 0.5440	obs per group: min	=	107
between = 0.9989	avg	=	107.0
overall = 0.9520	max	=	107
corr(u_i, Xb) = 0.9451	F(3, 3)	=	.
	Prob > F	=	.

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.483284	.024683	19.58	0.000	.4047317	.5618362
x4	.803612	.3188837	2.52	0.086	-.2112182	1.818442
x9	.0089608	.0011252	7.96	0.004	.00538	.0125416
x10	.0075775	.0030303	2.50	0.088	-.0020662	.0172212
x15	.3612625	.201384	1.79	0.171	-.2796314	1.002156
_cons	-1.742882	.1687559	-10.33	0.002	-2.279938	-1.205825
sigma_u	.33765923					
sigma_e	.09484762					
rho	.92686706	(fraction of variance due to u_i)				

## 6.4.- ESTIMACIÓN DEL MODELO DINÁMICO DE EFECTOS INDIVIDUALES ESPECÍFICOS FORMATO DE 900 A 1199 g.

La primera aproximación al modelo es:

```
. xtreg logproporcion retardo1logproporcion retardo2logproporcion retardo3
> logproporcion retardo4logproporcion x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12
> x13 x14 x15 x16 x17 x18 x19 x20 x21 x22 x23 x24 x25 x26 x27 x28 x29 x30 x31
> x32 x33 x34 x35, fe vce(cluster panel )
note: x26 omitted because of collinearity
note: x33 omitted because of collinearity
```

```
Fixed-effects (within) regression              Number of obs   =       416
Group variable: panel                        Number of groups =        4

R-sq:    within  = 0.6970                    Obs per group:  min =       104
          between = 0.9800                      avg   =     104.0
          overall  = 0.6434                      max   =       104

corr(u_i, xb) = 0.3452                      F(3,3)          =      .
                                          Prob > F         =      .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.4211132	.0886376	4.75	0.018	.1390287	.7031977
retardo2lo~n	-.1140621	.0792222	-1.44	0.246	-.3661818	.1380576
retardo3lo~n	.1154994	.0707995	1.63	0.201	-.1098163	.340815
retardo4lo~n	-.006964	.0315903	-0.22	0.840	-.1074983	.0935704
x1	-.0030153	.0009166	-3.29	0.046	-.0059323	-.0000982
x2	.6283297	.1500197	4.19	0.025	.1509	1.105759
x3	.1703034	.0327765	5.20	0.014	.0659938	.274613
x4	.3548472	.2869971	1.24	0.304	-.5585058	1.2682
x5	.0014657	.0007081	2.07	0.130	-.0007879	.0037193
x6	.0003543	.0007437	0.48	0.666	-.0020125	.0027211
x7	-.002018	.0024463	-0.82	0.470	-.0098032	.0057672
x8	.0419812	.5734857	0.07	0.946	-1.783106	1.867069
x9	.0076297	.0021372	3.57	0.038	.0008282	.0144313
x10	.0057371	.0019388	2.96	0.060	-.000433	.0119072
x11	.0021175	.0024161	0.88	0.445	-.0055717	.0098067
x12	-.0018646	.0013705	-1.36	0.267	-.0062262	.002497
x13	.0039817	.0010752	3.70	0.034	.0005599	.0074035
x14	.0082245	.0031076	2.65	0.077	-.0016653	.0181142
x15	-.120192	.2417293	-0.50	0.653	-.8894824	.6490984
x16	-.0006654	.000769	-0.87	0.451	-.0031127	.0017819
x17	.0011779	.0001357	8.68	0.003	.000746	.0016098
x18	.0012523	.0006401	1.96	0.145	-.0007846	.0032893
x19	.0006398	.001504	0.43	0.699	-.0041467	.0054262
x20	.0012698	.0003942	3.22	0.049	.0000153	.0025244
x21	.0015141	.0013323	1.14	0.338	-.0027259	.005754
x22	.0735135	.0699694	1.05	0.371	-.1491604	.2961873
x23	.0011539	.0034017	0.34	0.757	-.0096718	.0119796
x24	.0008153	.003033	0.27	0.806	-.0088369	.0104675
x25	-.0003197	.000743	-0.43	0.696	-.0026841	.0020448
x26	(omitted)					
x27	.0009063	.0003327	2.72	0.072	-.0001526	.0019652
x28	.0365412	.0174868	2.09	0.128	-.0191095	.092192
x29	.0058658	.2035256	0.03	0.979	-.6418436	.6535752
x30	-.0016133	.0045936	-0.35	0.749	-.0162323	.0130057
x31	.0040152	.0020814	1.93	0.149	-.0026089	.0106392
x32	.0001553	.0006445	0.24	0.825	-.0018959	.0022065
x33	(omitted)					
x34	.0001317	.0006731	0.20	0.857	-.0020105	.0022739
x35	-.0600896	.0156867	-3.83	0.031	-.1100117	-.0101674
_cons	-1.803867	.6134489	-2.94	0.060	-3.756135	.1484012
sigma_u	.15875465					
sigma_e	.11872381					
rho	.64132509	(fraction of variance due to u_i)				

Ahora eliminaremos las variables no significativas para disponer de un modelo más simple y manejable.

La propuesta de eliminación de variables está recogida en el modelo siguiente.

```
. xtreg logproporcion retardo1logproporcion x1 x2 x9 x13 x14 x17 x18 x28 x35
> , fe vce(cluster panel)
```

```
Fixed-effects (within) regression           Number of obs   =       428
Group variable: panel                      Number of groups =        4

R-sq:  within = 0.6354                     Obs per group:  min =       107
          between = 0.9657                  avg   =      107.0
          overall = 0.7294                  max   =       107

corr(u_i, Xb) = 0.5061                     F(3,3)          =          .
                                          Prob > F        =          .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.4074064	.0745708	5.46	0.012	.1700888	.6447241
x1	-.0030689	.0003072	-9.99	0.002	-.0040467	-.0020911
x2	.5025073	.06791	7.40	0.005	.2863874	.7186272
x9	.0097239	.0027533	3.53	0.039	.0009615	.0184862
x13	.00677	.0007904	8.57	0.003	.0042547	.0092853
x14	.0086438	.0018395	4.70	0.018	.0027897	.014498
x17	.0013495	.0003664	3.68	0.035	.0001833	.0025156
x18	.0019352	.0002895	6.68	0.007	.0010137	.0028566
x28	.0505914	.0025134	20.13	0.000	.0425926	.0585902
x35	-.0405316	.0046653	-8.69	0.003	-.0553786	-.0256845
_cons	-1.463628	.1816437	-8.06	0.004	-2.0417	-.8855571
sigma_u	.10923176					
sigma_e	.12583835					
rho	.42970555	(fraction of variance due to u_i)				

A continuación realizamos un contraste para testar la significatividad conjunta de un subconjunto de parámetros. Los resultados del contraste nos indican que debemos **NO RECHAZAR** la hipótesis nula; que los parámetros asociados a las variables explicativas son cero. Por lo tanto es posible eliminar estas variables del modelo, para así obtener un modelo más sencillo y fácil de manejar.

Las variables eliminadas y el resultado del contraste se encuentran en el cuadro siguiente.

```

. testparm retardo2logproporcionN retardo3logproporcionN retardo4logproporcionN
> x3 x4 x5 x6 x7 x8 x10 x11 x12 x15 x16 x19 x20 x21 x22 x23 x24 x25 x26 x27 x2
> 9 x30 x31 x32 x33 x34

( 1)  retardo2logproporcionN = 0
( 2)  retardo3logproporcionN = 0
( 3)  retardo4logproporcionN = 0
( 4)  x3 = 0
( 5)  x4 = 0
( 6)  x5 = 0
( 7)  x6 = 0
( 8)  x7 = 0
( 9)  x8 = 0
(10)  x10 = 0
(11)  x11 = 0
(12)  x12 = 0
(13)  x15 = 0
(14)  x16 = 0
(15)  x19 = 0
(16)  x20 = 0
(17)  x21 = 0
(18)  x22 = 0
(19)  x23 = 0
(20)  x24 = 0
(21)  x25 = 0
(22)  x27 = 0
(23)  x30 = 0
(24)  x31 = 0
      Constraint 1 dropped
      Constraint 2 dropped
      Constraint 3 dropped
      Constraint 4 dropped
      Constraint 5 dropped
      Constraint 6 dropped
      Constraint 7 dropped
      Constraint 8 dropped
      Constraint 10 dropped
      Constraint 11 dropped
      Constraint 12 dropped
      Constraint 14 dropped
      Constraint 15 dropped
      Constraint 16 dropped
      Constraint 17 dropped
      Constraint 19 dropped
      Constraint 20 dropped
      Constraint 21 dropped
      Constraint 22 dropped
      Constraint 23 dropped
      Constraint 24 dropped

      F( 3, 3) = 0.92
      Prob > F = 0.5262

```

El modelo definitivo en este formato está compuesto por las siguientes variables:

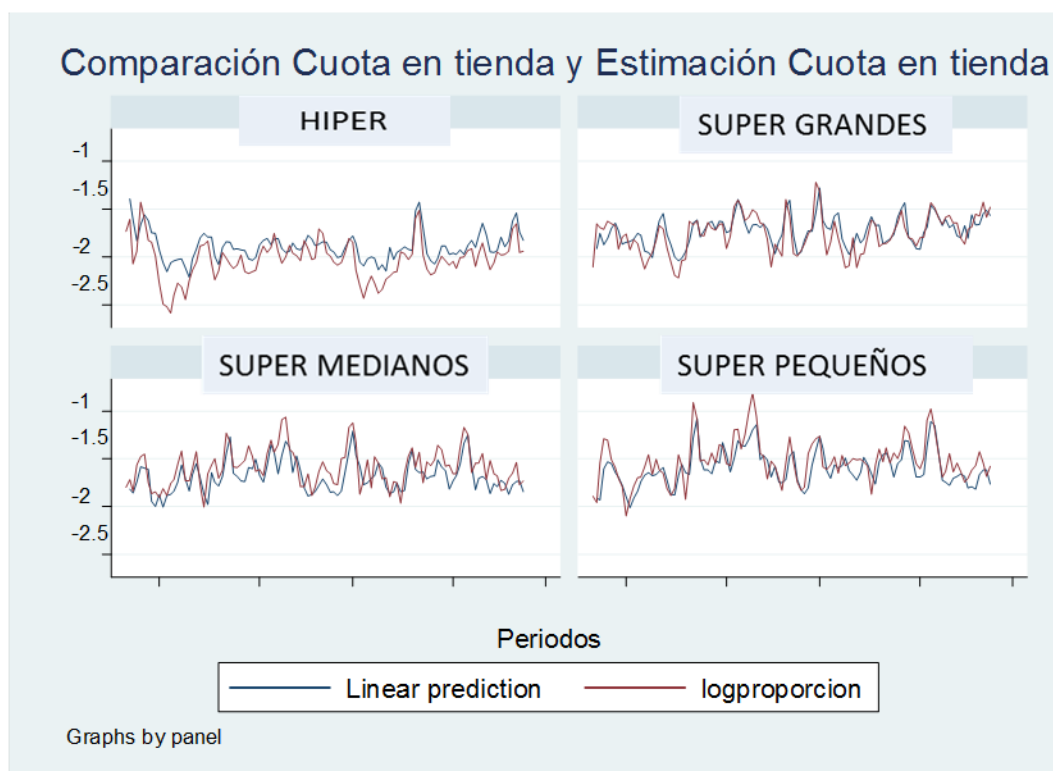
1. **Retardo1** Esta variable mide la influencia que tiene la variable a explicar (cuota de mercado) del periodo anterior. El signo asociado es positivo de tal forma que un aumento de la cuota de mercado en el periodo anterior produce un aumento en la variable a explicar del periodo actual (cuota de mercado).
2. **Distribución ponderada Marca 1:**  $x_1$  El signo asociado es negativo, cuando lo esperado es que fuera positivo. La explicación de esta situación sólo puede ser desde el punto de vista de marketing. Cabe pensar que si estamos presentes en más y mejores establecimientos, que es lo que mide la evolución positiva de esta variable, las ventas deberían evolucionar positivamente. Ahora bien, estamos modelizando la cuota en tienda, por lo que también es posible que a medida que mejoramos la distribución, nuestras ventas evolucionen positivamente pero nuestra cuota caiga, ya que estamos presentes en establecimientos que antes no

estábamos y lógicamente la participación de nuestra marca debe ser menor que en los que está presente desde tiempo atrás.

3. **Diferencial de precios respecto al Principal Competidor:  $x_2$**  El signo asociado a la variable es positivo y cumple con lo esperado. Cuanto mayor es nuestro diferencial de precio, mayor debe ser nuestra cuota en tienda.
4. **Cobertura promoción Folleto y Exposición Especial Marca 1:  $x_9$**  El signo esperado de esta variable es positivo, y efectivamente es el que resulta en el modelo. La variable mide la influencia que representa esta acción de marketing realizada por la Marca 1 sobre su cuota en tienda.
5. **Cobertura Reducción Temporal de Precio de la Marca 1:  $x_{13}$**  Esta variable indica en qué grado o con qué intensidad se ha realizado la promoción de Reducción de Precios de la Marca 1. Presenta signo positivo que es el esperado, de manera que la utilización de esta acción en numerosos establecimientos provoca una evolución positiva de la cuota en tienda de la Marca 1.
6. **Cobertura Regalo Directo de la Marca 1:  $x_{14}$**  Esta variable mide la influencia de la Promoción de Regalo Directo sobre la cuota en tienda de la Marca 1. Valores elevados en ella están asociados con valores elevados de la variable y, moviéndose ambas variables en el mismo sentido, por lo que el signo asociado es positivo, tal y como aparece en la salida del programa.
7. **Cobertura Promoción sólo Exposición Especial Principal Competidor:  $x_{17}$**  El signo esperado es negativo, sin embargo la estimación proporciona un signo positivo. La posible explicación, desde la óptica de marketing, es que este tipo de acción realizada por la competencia, llama la atención del consumidor sobre el producto, pero cuando llega el momento de materializar la compra, no se realiza sobre el producto de la competencia sino sobre el producto de la Marca 1. Por lo tanto, estas acciones realizadas por la competencia tienen un efecto positivo sobre la cuota en tienda de la Marca 1.

8. **Cobertura Promoción sólo Folleto Principal Competidor:  $x_{18}$**  La variable mide la influencia de un tipo especial de promoción realizada por el principal competidor. Parece lógico esperar un signo negativo, indicando una relación inversa entre las variables; ahora bien, en la muestra aparece signo positivo (relación directa entre las variables). Esta circunstancia puede interpretarse, desde el punto de vista de marketing, como que una mayor actividad promocional por parte de la competencia llama la atención sobre la categoría de producto, logrando que todas las marcas se beneficien.
9. **Cobertura Regalo Directo Resto Fabricantes:  $x_{28}$**  Esta variable nos indica, de nuevo, la influencia de un tipo especial de promoción (regalo directo) realizada en este caso por el Resto de Fabricantes. La interpretación que tiene es exactamente la misma que la variable  $x_{18}$ , incluso en su contradicción aparente.
10. **Cobertura Regalo Directo Marcas de Distribución:  $x_{35}$**  En esta ocasión la variable indica que una acción de Regalo Directo llevada a cabo por la competencia que representan las Marcas de Distribución produce un efecto negativo en la cuota de mercado de la marca. El signo es totalmente lógico y coincide con lo esperado.

En el gráfico se observa la comparación entre las cuotas de venta estimadas y reales en los 4 tipos de establecimientos considerados.



Aunque los establecimientos se han tratado como si sus efectos fueran iguales; esta es la hipótesis de construcción de los modelos de datos de panel, aun así vamos a comprobar si efectivamente las variables tienen el mismo efecto en los distintos tipos de establecimiento o si por el contrario, los efectos son diferentes.

Para comprobarlo volvemos a la estimación con variables ficticias por tipo de establecimiento. Los resultados de dicha estimación se muestran a continuación:

```
. regress logproporcion retardo1logproporcion x1 x2 x9 x13 x14 x17 x18 x28 x35
> r1h r1SG r1SM intx1hiper intx1SG intx1SM intx2hiper intx2SG intx2SM intx9hip
> er intx9SG intx9SM intx13hiper intx13SG intx13SM intx14hiper intx14SG intx14S
> M intx17hiper intx17SG intx17SM intx18hiper intx18SG intx18SM hiper superG su
> perM, vce(cluster panel)
```

Linear regression

Number of obs = 428  
F( 0, 3) = .  
Prob > F = .  
R-squared = 0.8259  
Root MSE = .12456

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.2710677	1.97e-12	1.4e+11	0.000	.2710677	.2710677
x1	-.0033366	1.77e-14	-1.9e+11	0.000	-.0033366	-.0033366
x2	.6882788	1.02e-11	6.7e+10	0.000	.6882788	.6882788
x9	.0133875	1.68e-14	7.9e+11	0.000	.0133875	.0133875
x13	.0076079	1.96e-15	3.9e+12	0.000	.0076079	.0076079
x14	.0207448	1.89e-13	1.1e+11	0.000	.0207448	.0207448
x17	-.0007436	2.40e-14	-3.1e+10	0.000	-.0007436	-.0007436
x18	.0015994	1.66e-14	9.7e+10	0.000	.0015994	.0015994
x28	.0287765	3.02e-12	9.5e+09	0.000	.0287765	.0287765
x35	-.0398138	8.86e-13	-4.5e+10	0.000	-.0398138	-.0398138
r1h	.3028806	4.73e-12	6.4e+10	0.000	.3028806	.3028806
r1SG	.1940118	2.15e-12	9.0e+10	0.000	.1940118	.1940118
r1SM	-.0157004	2.05e-12	-7.7e+09	0.000	-.0157004	-.0157004
intx1hiper	-.0137444	2.38e-12	-5.8e+09	0.000	-.0137444	-.0137444
intx1SG	.000204	2.19e-14	9.3e+09	0.000	.000204	.000204
intx1SM	-.0004315	2.01e-14	-2.1e+10	0.000	-.0004315	-.0004315
intx2hiper	-.3100004	1.21e-11	-2.6e+10	0.000	-.3100004	-.3100004
intx2SG	-.2031024	1.02e-11	-2.0e+10	0.000	-.2031024	-.2031024
intx2SM	-.0125798	1.03e-11	-1.2e+09	0.000	-.0125798	-.0125798
intx9hiper	.0015238	2.43e-13	6.3e+09	0.000	.0015238	.0015238
intx9SG	-.0098989	2.27e-14	-4.4e+11	0.000	-.0098989	-.0098989
intx9SM	-.0016196	1.84e-14	-8.8e+10	0.000	-.0016196	-.0016196
intx13hiper	-.0020285	3.86e-14	-5.3e+10	0.000	-.0020285	-.0020285
intx13SG	.0005714	3.75e-15	1.5e+11	0.000	.0005714	.0005714
intx13SM	-.0014097	2.29e-15	-6.2e+11	0.000	-.0014097	-.0014097
intx14hiper	-.0189109	3.65e-13	-5.2e+10	0.000	-.0189109	-.0189109
intx14SG	-.0135657	1.90e-13	-7.1e+10	0.000	-.0135657	-.0135657
intx14SM	-.016347	1.89e-13	-8.7e+10	0.000	-.016347	-.016347
intx17hiper	.0003896	2.54e-14	1.5e+10	0.000	.0003896	.0003896
intx17SG	.0052297	2.52e-14	2.1e+11	0.000	.0052297	.0052297
intx17SM	.0024995	2.43e-14	1.0e+11	0.000	.0024995	.0024995
intx18hiper	.000931	2.08e-14	4.5e+10	0.000	.000931	.000931
intx18SG	.000182	1.67e-14	1.1e+10	0.000	.000182	.000182
intx18SM	-.0004713	1.67e-14	-2.8e+10	0.000	-.0004713	-.0004713
hiper	2.072192	2.52e-10	8.2e+09	0.000	2.072192	2.072192
superG	.4311997	1.49e-11	2.9e+10	0.000	.4311997	.4311997
superM	.0726344	1.47e-11	4.9e+09	0.000	.0726344	.0726344
_cons	-1.797118	1.47e-11	-1.2e+11	0.000	-1.797118	-1.797118

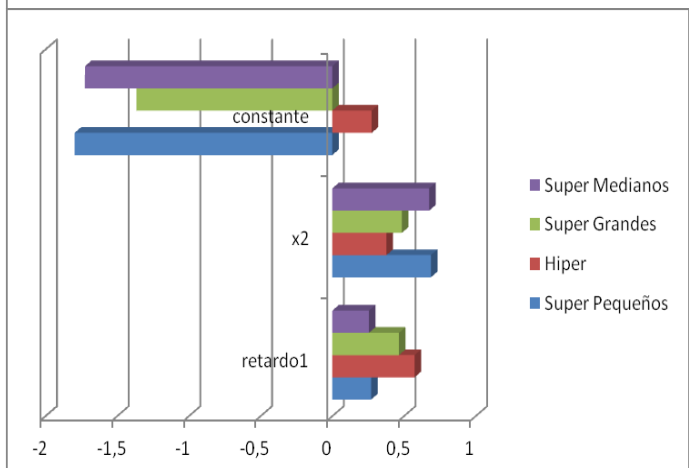
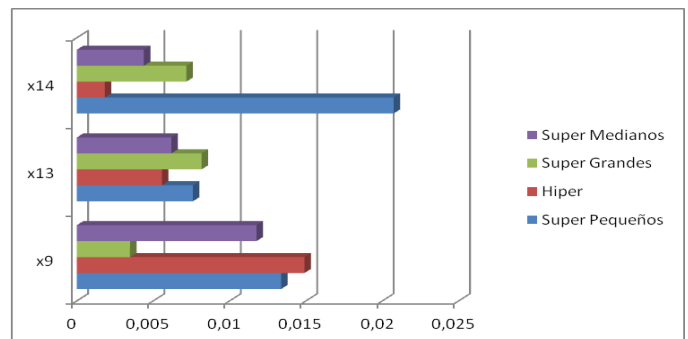
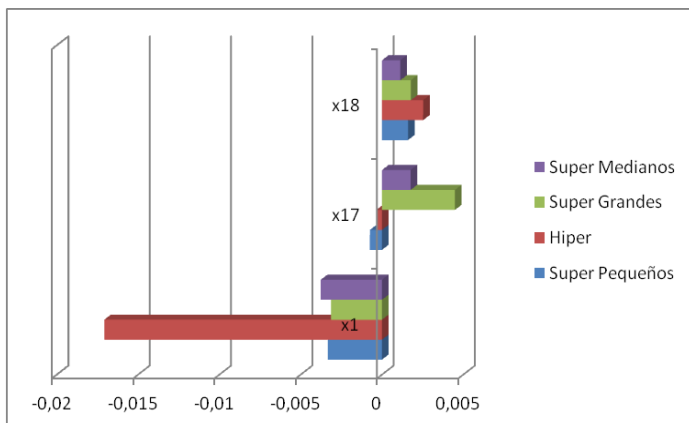
Las variables que nos interesan en este modelo son las ficticias híper, superG y superM, que se sumarán o restarán a la constante general del modelo según su signo y recogerán la influencia del tipo de establecimiento. También son objeto de interés las variables generadas como interacción entre cada una de las variables significativas y las ficticias anteriores. El coeficiente de cada una de ellas muestra el efecto diferencial de la variable considerada en el tipo de establecimiento.

Los resultados se muestran en el cuadro siguiente, donde se aprecian las diferencias que se producen en los distintos tipos de establecimiento.



	Super Pequeños	Hiper	Super Grandes	Super Medianos
<i>retardo1</i>	0,2710677	0,5739483	0,4650795	0,2553673
$x_1$	-0,0033366	-0,017081	-0,0031326	-0,0037681
$x_2$	0,6882788	0,3782784	0,4851764	0,675699
$x_9$	0,0133875	0,0149113	0,0034886	0,0117679
$x_{13}$	0,0076079	0,0055794	0,0081793	0,0061982
$x_{14}$	0,0207448	0,0018339	0,0071791	0,0043978
$x_{17}$	-0,0007436	-0,000354	0,0044861	0,0017559
$x_{18}$	0,0015994	0,0025304	0,0017814	0,0011281
$x_{28}$	0,0287765	0,0287765	0,0287765	0,0287765
$x_{35}$	-0,0398138	-0,0398138	-0,0398138	-0,0398138
<i>constante</i>	-1,797118	0,275074	-1,3659183	-1,7244836

En las variables  $x_{28}$  y  $x_{35}$  no se recogen los efectos individuales pues han sido eliminadas del análisis al producirse un problema de colinealidad.



En estos gráficos se muestran de una forma más visual las diferencias de efecto que presentan las variables en los distintos tipos de establecimiento.

El último paso es comprobar si el modelo está correctamente especificado; para cubrir este objetivo planteamos el contraste Reset.

Los resultados del contraste son los siguientes:

```
. xtreg logproporcion retardollogproporcion x1 x2 x9 x13 x14 x17 x18 x28 x35 prediccionfe
> 2, fe vce(cluster panel )
```

Fixed-effects (within) regression  
Group variable: panel

Number of obs = 428  
Number of groups = 4

R-sq: within = 0.6367  
between = 0.9667  
overall = 0.7341

Obs per group: min = 107  
avg = 107.0  
max = 107

corr(u\_i, xb) = 0.5118

F(3,3) = .  
Prob > F = .

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
retardollo~n	.1945505	.2529319	0.77	0.498	-.6103917 .9994928
x1	-.0015177	.0024596	-0.62	0.581	-.0093453 .0063099
x2	.2418184	.4372334	0.55	0.619	-1.149653 1.63329
x9	.0050562	.0066128	0.76	0.500	-.0159886 .026101
x13	.0033563	.0055478	0.60	0.588	-.0142994 .0210119
x14	.0037673	.0087262	0.43	0.695	-.0240035 .0315381
x17	.0005613	.0013001	0.43	0.695	-.0035763 .0046988
x18	.0009419	.0012378	0.76	0.502	-.0029973 .004881
x28	.0225967	.0426195	0.53	0.633	-.1130375 .1582309
x35	-.0154704	.0387337	-0.40	0.716	-.1387381 .1077974
prediccion~2	-.1535997	.2184844	-0.70	0.533	-.8489145 .5417152
_cons	-1.137938	.6196085	-1.84	0.164	-3.109809 .8339323
sigma_u	.10691069				
sigma_e	.12575899				
rho	.41951928	(fraction of variance due to u_i)			

```
. testparm prediccionfe2
```

( 1) prediccionfe2 = 0

F( 1, 3) = 0.49  
Prob > F = 0.5327

En el contraste paramétrico se observa que se acepta la hipótesis nula de que la variable predicción al cuadrado es no significativa. Estamos en el buen camino ya que no indica que hasta el momento la especificación realizada es la adecuada.

Sin embargo cuando en el contraste se incluyen las variables predicción al cuadrado y predicción al cubo la situación cambia. En el contraste paramétrico que sigue puede observarse que no es posible aceptar la hipótesis de que ambas variables no son significativas, pues el p valor asociado a ellas es 0,049 y que es inferior al nivel de significación del 5 %. Como consecuencia de este resultado, la forma funcional empleada no es correcta, esto es, existe alguna variable, o alguna interacción entre

variables que no se ha considerado y el contraste Reset la está poniendo de manifiesto.

```

. xreg logproportion retardo1logproportion x1 x2 x9 x13 x14 x17 x18 x28 x35 prediccionfe
> 2 prediccionfe3, fe vce(cluster panel)

Fixed-effects (within) regression                               Number of obs   =       428
Group variable: panel                                         Number of groups   =        4

R-sq:  within = 0.6429                                         Obs per group: min =       107
                   between = 0.9661                             avg       =      107.0
                   overall  = 0.7364                             max       =      107

                                F(3,3) = .
                                Prob > F   = .

corr(u_i, xb) = 0.5086

                                (Std. Err. adjusted for 4 clusters in panel)

+-----+-----+-----+-----+-----+-----+
| logproportion |      Coef.   | Robust Std. Err. |      t      | P>|t| | [95% Conf. Interval] |
+-----+-----+-----+-----+-----+-----+
| retardo1logproportion | 3.536984 | .6828879 | 5.18 | 0.014 | 1.36373 | 5.710238 |
| x1 | -0.026858 | .0055841 | -4.81 | 0.017 | -.0446267 | -.0090849 |
| x2 | 4.380036 | .9267074 | 4.73 | 0.018 | 1.43084 | 7.329233 |
| x9 | .0857845 | .0204365 | 4.20 | 0.025 | .0207464 | .1508226 |
| x13 | .059289 | .0112691 | 5.26 | 0.013 | .0234257 | .0951523 |
| x14 | .0748166 | .0151553 | 4.94 | 0.016 | .0265857 | .1230475 |
| x17 | .0117255 | .0020776 | 5.64 | 0.011 | .0051137 | .0183374 |
| x18 | .0168038 | .00348 | 4.83 | 0.017 | .0057288 | .0278787 |
| x28 | .4334425 | .0862912 | 5.02 | 0.015 | .1588255 | .7080594 |
| x35 | -.3395342 | .0703323 | -4.83 | 0.017 | -.5633629 | -.1157054 |
| prediccion2 | 4.92891 | 1.136206 | 4.34 | 0.023 | 1.312994 | 8.544825 |
| prediccion3 | 1.029754 | .2399254 | 4.29 | 0.023 | .2662047 | 1.793304 |
| _cons | -8.838879 | 1.778948 | -4.97 | 0.016 | -14.50029 | -3.177472 |
+-----+-----+-----+-----+-----+-----+
| sigma_u | .1073472 |
| sigma_e | .12483332 |
| rho | .42511218 | (fraction of variance due to u_i)
+-----+-----+-----+-----+-----+-----+

. testparm prediccionfe2 prediccionfe3
unrecognized command: testparm
r(199);

. testparm prediccionfe2 prediccionfe3

( 1) prediccionfe2 = 0
( 2) prediccionfe3 = 0

F( 2, 3) = 9.70
Prob > F = 0.0490

```

El problema ahora es determinar que variables han sido omitidas de la forma funcional. Como el problema aparece cuando se considera la variable predicción al cubo, las primeras candidatas a ser incluidas en la nueva forma funcional son las variables originales elevadas al cubo además de las que hasta ese momento están incluidas.

```

. xtreg logproportion retardollogproportion x1 x2 x9 x13 x14 x17 x18 x28 x35 x18cubo x35cubo
> bo, fe vce(cluster panel)

Fixed-effects (within) regression                               Number of obs   =       428
Group variable: panel                                         Number of groups   =        4

R-sq:  within = 0.6408                                         obs per group: min =       107
               between = 0.9630                                 avg               =      107.0
               overall  = 0.7090                                 max               =       107

               F(3,3)                                           =          .
               Prob > F                                           =          .

corr(u_i, xb) = 0.4725                                         F(3,3)           =          .
               Prob > F                                           =          .

               (Std. Err. adjusted for 4 clusters in panel)

+-----+-----+-----+-----+-----+-----+
| logproporc~n |          Coef.   | Robust Std. Err. |      t      | P>|t| | [95% Conf. Interval] |
+-----+-----+-----+-----+-----+-----+
| retardol1o~n | .3966162         | .0756255         |    5.24     | 0.013 | .1559421   .6372904   |
| x1            | -.0027039        | .0002701         |   -10.01    | 0.002 | -.0035636  -.0018443   |
| x2            | .5402461         | .0712343         |    7.58     | 0.005 | .3135468   .7669453   |
| x9            | .0098596         | .0026055         |    3.78     | 0.032 | .0015677   .0181515   |
| x13           | .0065626         | .0007806         |    8.41     | 0.004 | .0040785   .0090468   |
| x14           | .0089146         | .0019022         |    4.69     | 0.018 | .0028609   .0149683   |
| x17           | .0013155         | .000385         |    3.42     | 0.042 | .0000902   .0025408   |
| x18           | .0046501         | .0005196         |    8.95     | 0.003 | .0029963   .0063038   |
| x28           | .0547714         | .0029134         |   18.80     | 0.000 | .0454996   .0640433   |
| x35           | .4758598         | .1988449         |    2.39     | 0.096 | -.1569536   1.108673   |
| x18cubo       | -2.50e-06        | 6.09e-07         |   -4.11     | 0.026 | -4.44e-06   -5.63e-07   |
| x35cubo       | -.1026378        | .0386774         |   -2.65     | 0.077 | -.2257264   .0204508   |
| _cons         | -1.552604        | .188162         |   -8.25     | 0.004 | -2.15142   -.9537889   |
+-----+-----+-----+-----+-----+-----+
| sigma_u       | .12109409        |                   |             |       |                   |
| sigma_e       | .12519967        |                   |             |       |                   |
| rho           | .48333517        |                   |             |       |                   |
|               | (fraction of variance due to u_i) |

```

Así se ha hecho, resultando que las únicas variables significativas son las variables x18 al cubo y x35 al cubo, las restantes se han eliminado ya que no eran significativas.

¿Será esta la forma funcional definitiva? Para comprobarlo tenemos que calcular de nuevo el contraste Reset, pero con estas nuevas variables.

```
. xtreg logproporción retardo1logproporción x1 x2 x9 x13 x14 x17 x18 x28 x35 x18cubo x35cu
> bo predicciónfex2, fe vce(cluster panel)
```

Fixed-effects (within) regression

Number of obs = 428

Group variable: panel

Number of groups = 4

R-sq: within = 0.6418

between = 0.9637

overall = 0.7136

obs per group: min = 107

avg = 107.0

max = 107

corr(u\_i, xb) = 0.4781

F(3,3) = .

Prob > F = .

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
retardo1lo~n	.2091572	.281621	0.74	0.512	-.6870866 1.105401
x1	-.0014719	.0023856	-0.62	0.581	-.009064 .0061203
x2	.2851197	.5119342	0.56	0.616	-1.344083 1.914323
x9	.0055521	.0074565	0.74	0.511	-.0181777 .0292818
x13	.0035692	.0060264	0.59	0.595	-.0156096 .0227481
x14	.0043876	.0096978	0.45	0.682	-.026475 .0352503
x17	.0006236	.0013327	0.47	0.672	-.0036176 .0048647
x18	.0023699	.0044971	0.53	0.635	-.0119418 .0166816
x28	.0272905	.0523042	0.52	0.638	-.1391647 .1937458
x35	.3498507	.2551522	1.37	0.264	-.4621575 1.161859
x18cubo	-1.18e-06	2.78e-06	-0.43	0.698	-.00001 7.65e-06
x35cubo	-.0730803	.0572009	-1.28	0.291	-.2551193 .1089586
prediccio~X2	-.139133	.2482364	-0.56	0.614	-.9291321 .650866
_cons	-1.214087	.7603417	-1.60	0.209	-3.633834 1.205659
sigma_u	.11911453				
sigma_e	.12517772				
rho	.47519589				(fraction of variance due to u_i)

```
. testparm predicciónfex2
```

( 1) predicciónfex2 = 0

F( 1, 3) = 0.31

Prob > F = 0.6143

La primera parte del contraste proporciona buenas noticias, no existe ninguna variable al cuadrado omitida, vamos con la segunda parte.

```
. xtreg logproporcion retardo1logproporcion x1 x2 x9 x13 x14 x17 x18 x28 x35 x18cubo x35cu
> bo prediccionfe2 prediccionfe3, fe vce(cluster panel )
```

```
Fixed-effects (within) regression                               Number of obs   =       428
Group variable: panel                                         Number of groups  =        4

R-sq:  within = 0.6469                                         Obs per group:   min =       107
        between = 0.9635                                         avg =      107.0
        overall = 0.7151                                         max =       107
```

```
corr(u_i, xb) = 0.4753                                         F(3,3) = .
Prob > F = .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	3.422803	.9474813	3.61	0.036	.407495	6.438112
x1	-.0235119	.0067669	-3.47	0.040	-.0450474	-.0019765
x2	4.67728	1.359534	3.44	0.041	.3506351	9.003924
x9	.0863071	.0273849	3.15	0.051	-.0008438	.173458
x13	.057125	.015383	3.71	0.034	.0081696	.1060805
x14	.0767605	.021842	3.51	0.039	.0072494	.1462716
x17	.0113792	.0028883	3.94	0.029	.0021874	.020571
x18	.0399799	.0112094	3.57	0.038	.0043067	.0756531
x28	.4671864	.1297449	3.60	0.037	.0542802	.8800927
x35	4.55872	1.111581	4.10	0.026	1.021172	8.096267
x18cubo	-.0000213	6.12e-06	-3.48	0.040	-.0000408	-1.79e-06
x35cubo	-.9733704	.2393813	-4.07	0.027	-1.735189	-.2115523
prediccion~x2	4.871483	1.589082	3.07	0.055	-.1856864	9.928652
prediccion~x3	1.014446	.3381158	3.00	0.058	-.061589	2.090482
_cons	-9.52901	2.636955	-3.61	0.036	-17.92098	-1.137043
sigma_u	.11958314					
sigma_e	.12444486					
rho	.48008509	(fraction of variance due to u_i)				

```
. testparm prediccionfe2 prediccionfe3
variable prediccionfe3 not found
r(111);
```

```
. testparm prediccionfe2 prediccionfe3
```

```
( 1) prediccionfe2 = 0
( 2) prediccionfe3 = 0

F( 2, 3) = 5.37
Prob > F = 0.1021
```

Esta vez, el contraste nos dice que la forma funcional utilizada es correcta ya que se rechaza la hipótesis de que las variables predicción al cuadrado y al cubo son significativas. Por tanto, a partir de aquí obtendremos el modelo definitivo para el formato de 900 a 1199 gr.

El modelo definitivo quitando las variables que se han utilizado para el contraste Reset es:

```
. xtreg logproporcion retardo1logproporcion x1 x2 x9 x13 x14 x17 x18 x28 x35 x18cubo x35cu
> bo, fe vce(cluster panel )
```

```
Fixed-effects (within) regression                               Number of obs   =       428
Group variable: panel                                         Number of groups  =        4
```

```
R-sq:  within = 0.6408                                         Obs per group:   min =       107
        between = 0.9630                                         avg =      107.0
        overall = 0.7090                                         max =       107
```

```
corr(u_i, xb) = 0.4725                                         F(3,3) = .
Prob > F = .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.3966162	.0756255	5.24	0.013	.1559421	.6372904
x1	-.0027039	.0002701	-10.01	0.002	-.0035636	-.0018443
x2	.5402461	.0712343	7.58	0.005	.3135468	.7669453
x9	.0098596	.0026055	3.78	0.032	.0015677	.0181515
x13	.0065626	.0007806	8.41	0.004	.0040785	.0090468
x14	.0089146	.0019022	4.69	0.018	.0028609	.0149683
x17	.0013155	.000385	3.42	0.042	.0000902	.0025408
x18	.0046501	.0005196	8.95	0.003	.0029963	.0063038
x28	.0547714	.0029134	18.80	0.000	.0454996	.0640433
x35	.4758598	.1988449	2.39	0.096	-.1569536	1.108673
x18cubo	-2.50e-06	6.09e-07	-4.11	0.026	-4.44e-06	-5.63e-07
x35cubo	-.1026378	.0386774	-2.65	0.077	-.2257264	.0204508
_cons	-1.552604	.188162	-8.25	0.004	-2.15142	-.9537889
sigma_u	.12109409					
sigma_e	.12519967					
rho	.48333517	(fraction of variance due to u_i)				

## 6.5.- ESTIMACIÓN DEL MODELO DINÁMICO DE EFECTOS INDIVIDUALES ESPECÍFICOS FORMATO DE 1200 A 1499 g.

La primera aproximación al modelo es:

```
. xtreg logproporción retardo1logproporción retardo2logproporción retardo3logproporción re
> tardo4logproporción x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 x13 x14 x15 x16 x17 x18 x19 x20
> x21 x22 x23 x24 x25 x26 x27 x28 x29 x30 x31 x32 x33 x34 x35, fe vce (cluster panel)
note: x3 omitted because of collinearity
note: x6 omitted because of collinearity
note: x22 omitted because of collinearity
note: x23 omitted because of collinearity
note: x24 omitted because of collinearity
note: x25 omitted because of collinearity
note: x26 omitted because of collinearity
note: x27 omitted because of collinearity
note: x28 omitted because of collinearity
note: x33 omitted because of collinearity
note: x35 omitted because of collinearity

Fixed-effects (within) regression              Number of obs   =       416
Group variable: panel                          Number of groups =        4

R-sq:  within = 0.8298                        obs per group: min =       104
        between = 0.9961                      avg =       104.0
        overall = 0.9183                      max =       104

corr(u_i, Xb) = 0.7280                        F(3,3)          =        .
                                                Prob > F        =        .

                                (Std. Err. adjusted for 4 clusters in panel)

+-----+-----+-----+-----+-----+-----+
| logproporc~n | Coef. | Robust Std. Err. | t | P>|t| | [95% Conf. Interval] |
+-----+-----+-----+-----+-----+
| retardo1lo~n | .3711591 | .1272952 | 2.92 | 0.062 | -.0339509 | .7762691 | |
| retardo2lo~n | .0299735 | .0283184 | 1.06 | 0.368 | -.0601483 | .1200953 |
| retardo3lo~n | .0814371 | .0350211 | 2.33 | 0.103 | -.0300158 | .1928899 |
| retardo4lo~n | .1275855 | .0143059 | 8.92 | 0.003 | .0820577 | .1731133 |
| x1 | -.0014609 | .0011649 | -1.25 | 0.299 | -.005168 | .0022462 |
| x2 | 1.600718 | .602894 | 2.66 | 0.077 | -.3179599 | 3.519396 |
| x3 | (omitted) | | | | | | |
| x4 | -.0654312 | .2381633 | -0.27 | 0.801 | -.8233732 | .6925107 |
| x5 | -.0017043 | .0023516 | -0.72 | 0.521 | -.0091881 | .0057794 |
| x6 | (omitted) | | | | | | |
| x7 | .0029659 | .0014592 | 2.03 | 0.135 | -.0016779 | .0076097 |
| x8 | -.0102919 | .1491252 | -0.07 | 0.949 | -.4848749 | .464291 |
| x9 | .0111932 | .0033859 | 3.31 | 0.046 | .0004177 | .0219687 |
| x10 | .0045035 | .0024248 | 1.86 | 0.160 | -.0032133 | .0122203 |
| x11 | .0014474 | .0009978 | 1.45 | 0.243 | -.0017281 | .0046228 |
| x12 | -.0099793 | .0047751 | -2.09 | 0.128 | -.0251756 | .0052171 |
| x13 | .0048098 | .000635 | 7.57 | 0.005 | .0027889 | .0068307 |
| x14 | -.0034534 | .0035582 | -0.97 | 0.403 | -.0147773 | .0078704 |
| x15 | -.0260145 | .136109 | -0.19 | 0.861 | -.459174 | .4071451 |
| x16 | .0088386 | .0038329 | 2.31 | 0.104 | -.0033593 | .0210364 |
| x17 | .0060741 | .0012109 | 5.02 | 0.015 | .0022206 | .0099277 |
| x18 | -.0001647 | .0010008 | -0.16 | 0.880 | -.0033498 | .0030204 |
| x19 | .0036795 | .0253033 | 0.15 | 0.894 | -.0768468 | .0842059 |
| x20 | .0018011 | .0014892 | 1.21 | 0.313 | -.0029382 | .0065404 |
| x21 | .0019217 | .0006943 | 2.77 | 0.070 | -.000288 | .0041314 |
| x22 | (omitted) | | | | | | |
| x23 | (omitted) | | | | | | |
| x24 | (omitted) | | | | | | |
| x25 | (omitted) | | | | | | |
| x26 | (omitted) | | | | | | |
| x27 | (omitted) | | | | | | |
| x28 | (omitted) | | | | | | |
| x29 | -.0232335 | .282908 | -0.08 | 0.940 | -.9235729 | .877106 |
| x30 | .0029932 | .0017194 | 1.74 | 0.180 | -.0024787 | .0084652 |
| x31 | .0009559 | .0008962 | 1.07 | 0.364 | -.0018963 | .003808 |
| x32 | -.0012999 | .0008948 | -1.45 | 0.242 | -.0041476 | .0015478 |
| x33 | (omitted) | | | | | | |
| x34 | -.000388 | .0006196 | -0.63 | 0.576 | -.0023599 | .0015839 |
| x35 | (omitted) | | | | | | |
| _cons | -2.505426 | .90105 | -2.78 | 0.069 | -5.372969 | .3621174 |
+-----+-----+-----+-----+-----+
| sigma_u | .17493342 | |
| sigma_e | .1633318 |
| rho | .53425712 | (fraction of variance due to u_i) |
+-----+-----+-----+-----+-----+

```

Resulta evidente la necesidad de eliminar las variables que producen colinealidad, así como aquellas variables que no son significativas.

El modelo que queda después de eliminar estas variables es el siguiente:

```
. xtreg logproporcion retardo1logproporcion retardo4logproporcion x9 x13 x16 x17, fe vce (c
> luster panel)
```

```
Fixed-effects (within) regression                Number of obs   =       416
Group variable: panel                          Number of groups =         4

R-sq:  within = 0.7997                        obs per group: min =       104
        between = 0.9996                      avg =       104.0
        overall = 0.9216                      max =       104

corr(u_i, xb) = 0.7779                        F(3,3)          =         .
                                                Prob > F         =         .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.5452319	.0642719	8.48	0.003	.34069	.7497738
retardo4lo~n	.2354443	.056325	4.18	0.025	.0561932	.4146955
x9	.0136576	.0019721	6.93	0.006	.0073815	.0199336
x13	.0075902	.0008369	9.07	0.003	.0049269	.0102534
x16	.0031292	.0008927	3.51	0.039	.0002881	.0059704
x17	.0034413	.0003504	9.82	0.002	.0023263	.0045563
_cons	-.5849709	.0347507	-16.83	0.000	-.6955633	-.4743785
sigma_u	.12991539					
sigma_e	.17229366					
rho	.36247596	(fraction of variance due to u_i)				

Como en los casos anteriores vamos a contrastar si las variables eliminadas son conjuntamente significativas.

```
. testparm retardo2logproporcion x1 x2 x4 x5 x6 x7 x8 x10 x11 x12 x1
> 4 x15 x18 x19 x20 x21 x29 x30 x31 x32 x34
```

```
( 1) retardo2logproporcion = 0
( 2) retardo3logproporcion = 0
( 3) x1 = 0
( 4) x2 = 0
( 5) x4 = 0
( 6) x5 = 0
( 7) x7 = 0
( 8) x8 = 0
( 9) x10 = 0
(10) x11 = 0
(11) x12 = 0
(12) x14 = 0
(13) x15 = 0
(14) x18 = 0
(15) x19 = 0
(16) x20 = 0
(17) x21 = 0
(18) x29 = 0
(19) x30 = 0
(20) x31 = 0
(21) x32 = 0
(22) x34 = 0
```

```
Constraint 1 dropped
Constraint 2 dropped
Constraint 3 dropped
Constraint 6 dropped
Constraint 7 dropped
Constraint 8 dropped
Constraint 9 dropped
Constraint 10 dropped
Constraint 11 dropped
Constraint 12 dropped
Constraint 13 dropped
Constraint 14 dropped
Constraint 15 dropped
Constraint 16 dropped
Constraint 17 dropped
Constraint 19 dropped
Constraint 20 dropped
Constraint 21 dropped
Constraint 22 dropped
```

```
F( 3, 3) = 4.83
Prob > F = 0.1141
```

Los resultados del contraste nos indican que podemos eliminar las variables del modelo porque no son conjuntamente significativas, y así trabajar con un modelo más sencillo y manejable.

El modelo definitivo está compuesto por las variables siguientes:

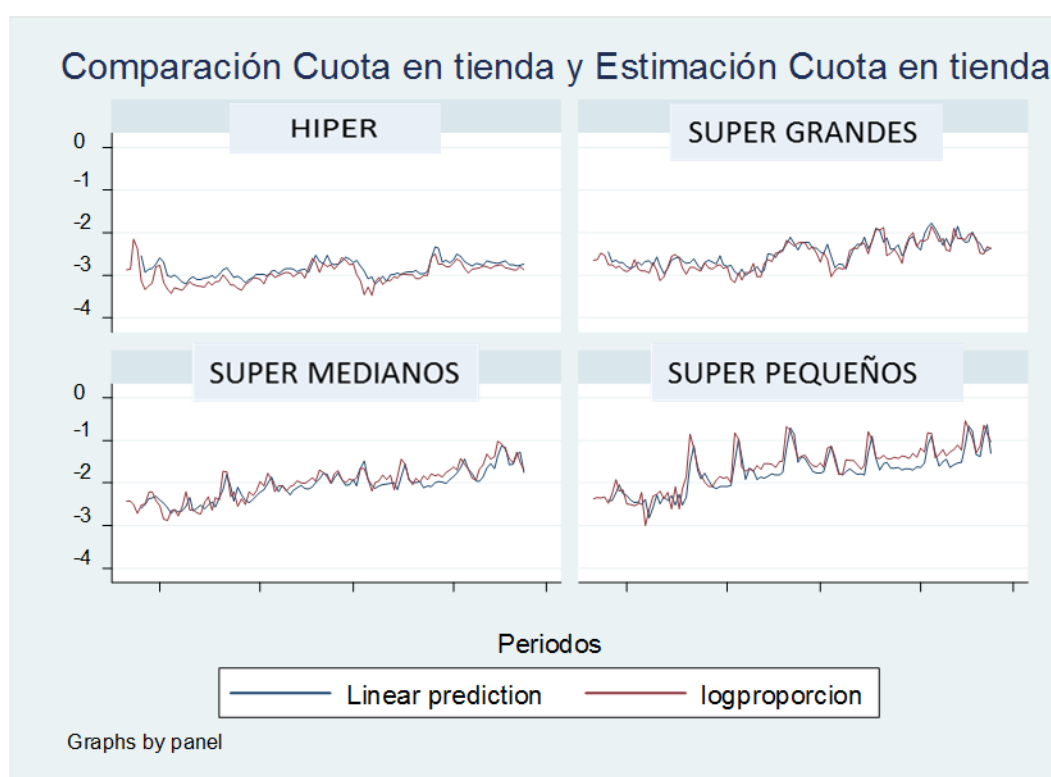
1. **Retardo1** Esta variable mide la influencia que tiene la variable cuota de mercado del periodo anterior. El signo asociado es positivo de tal forma que un aumento de la cuota de mercado en el periodo anterior produce un aumento en la variable a explicar del periodo actual.
2. **Retardo4** Mide la influencia de la variable cuota de mercado de 4 periodos atrás. El signo asociado es positivo de manera que la relación entre la variable a explicar y esta variable es directa; aumentos en una suponen aumentos en la otra y viceversa.
3. **Cobertura Promoción Folleto y Exposición Especial Marca 1:  $x_9$**  El signo esperado de esta variable es positivo, que efectivamente se confirma en el análisis. La variable mide el porcentaje de establecimientos en los que la Marca 1 realiza este tipo de promoción. Cuanto mayor es este valor mayor debe ser la cuota en tienda de la Marca 1.
4. **Cobertura Reducción Temporal de Precios de la Marca 1:  $x_{13}$**  Esta variable habla de la influencia, en este caso positiva, que supone realizar este tipo de promoción. Es de resaltar que lo importante no es la cuantía de la promoción, sino que la promoción sea muy visible; esto es, que esté en muchos establecimientos a la vez. Esto es lo que mide la variable cobertura.
5. **Cobertura Promoción Folleto y Exposición Especial Principal Competidor:  $x_{16}$**  Esta variable presenta un signo positivo cuando lo esperado según la teoría debería ser un signo negativo. La explicación puede venir desde el punto de vista de marketing. Este tipo de acciones llama la atención sobre toda la categoría de producto y cuando llega el momento de materializar la compra no se materializa sobre la marca que la realiza, el competidor principal, sino sobre la Marca 1.



## 6. Cobertura Promoción sólo Exposición Especial Principal Competidor: $x_{17}$

Nuevamente signo diferente al esperado ya que las acciones de la competencia deberían tener un efecto negativo sobre la Marca 1. La explicación no puede ser otra que la que hemos dado en el punto anterior. Este tipo de acciones llaman la atención sobre la categoría de producto de manera que la Marca 1 sale beneficiada como marca líder en el mercado.

A continuación, en el gráfico, vemos la comparación entre las cuotas de venta estimadas y reales en los 4 tipos de establecimientos considerados.



El paso siguiente es comprobar si los efectos de las variables independientes sobre la variable a explicar son los mismos según los diferentes tipos de establecimientos. Para ello, calculamos la regresión sobre la variable a explicar respecto a las variables que han resultado significativas en el modelo anterior, y a estas variables les añadimos: la interacción entre esas variables y los distintos tipos de establecimientos más las variables ficticias "tipos de establecimiento".

Los resultados de esta regresión son los siguientes:

```
. regress logproporcion retardo1logproporcion retardo4logproporcionN x9 x13 x16 x17 r1h r1s
> g r1sm r4h r4sg r4sm intx9hiper intx9superg intx9superm intx13hiper intx13superg intx13supe
> rm intx16hiper intx16superg intx16superm intx17hiper intx17superg intx17superm hiper superg
> superm, vce (cluster panel)
note: intx16hiper omitted because of collinearity
note: intx16superg omitted because of collinearity
note: intx16superm omitted because of collinearity
```

```
Linear regression      Number of obs =      416
                      F( 0, 3) =      .
                      Prob > F =      .
                      R-squared =    0.9358
                      Root MSE =    .17208
```

(Std. Err. adjusted for 4 clusters in panel)

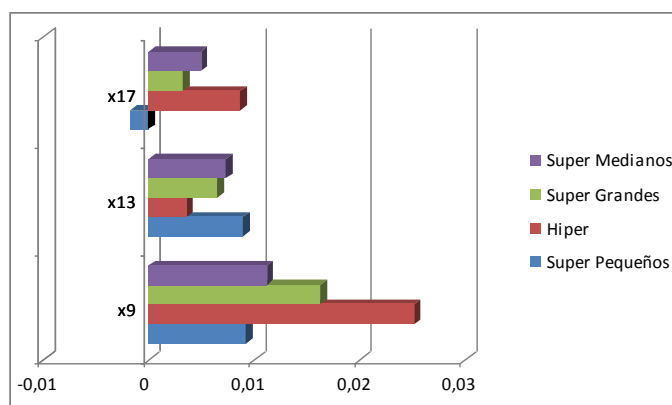
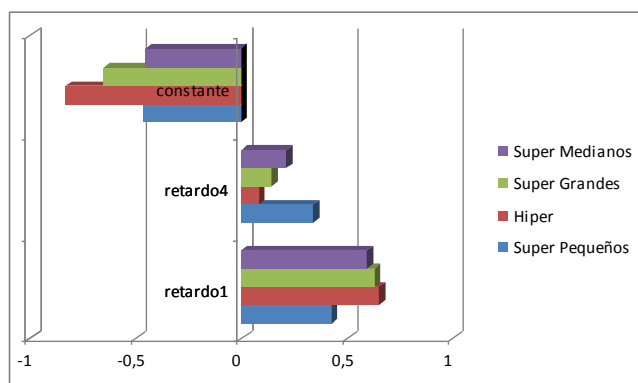
logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.4281685	3.04e-13	1.4e+12	0.000	.4281685	.4281685
retardo4lo~n	.340386	5.53e-13	6.2e+11	0.000	.340386	.340386
x9	.0092369	2.62e-14	3.5e+11	0.000	.0092369	.0092369
x13	.0090225	7.76e-17	1.2e+14	0.000	.0090225	.0090225
x16	.0015965	1.03e-14	1.5e+11	0.000	.0015965	.0015965
x17	-.0017983	1.08e-14	-1.7e+11	0.000	-.0017983	-.0017983
r1h	.2277703	1.12e-12	2.0e+11	0.000	.2277703	.2277703
r1sg	.2040719	4.52e-13	4.5e+11	0.000	.2040719	.2040719
r1sm	.1672583	3.19e-13	5.2e+11	0.000	.1672583	.1672583
r4h	-.2581496	1.68e-12	-1.5e+11	0.000	-.2581496	-.2581496
r4sg	-.1947659	5.54e-13	-3.5e+11	0.000	-.1947659	-.1947659
r4sm	-.1253518	5.55e-13	-2.3e+11	0.000	-.1253518	-.1253518
intx9hiper	.0160097	2.72e-14	5.9e+11	0.000	.0160097	.0160097
intx9superg	.0070999	2.73e-14	2.6e+11	0.000	.0070999	.0070999
intx9superm	.0019956	2.62e-14	7.6e+10	0.000	.0019956	.0019956
intx13hiper	-.0054185	1.23e-14	-4.4e+11	0.000	-.0054185	-.0054185
intx13superg	-.0025053	3.34e-15	-7.5e+11	0.000	-.0025053	-.0025053
intx13superm	-.0016575	1.74e-15	-9.5e+11	0.000	-.0016575	-.0016575
intx16hiper	(omitted)					
intx16superg	(omitted)					
intx16superm	(omitted)					
intx17hiper	.0104801	2.32e-14	4.5e+11	0.000	.0104801	.0104801
intx17superg	.0050431	1.08e-14	4.7e+11	0.000	.0050431	.0050431
intx17superm	.0067827	1.08e-14	6.3e+11	0.000	.0067827	.0067827
hiper	-.368969	1.49e-12	-2.5e+11	0.000	-.368969	-.368969
superg	-.1893034	1.09e-12	-1.7e+11	0.000	-.1893034	-.1893034
superm	.0085715	4.39e-13	2.0e+10	0.000	.0085715	.0085715
_cons	-.4656497	4.22e-13	-1.1e+12	0.000	-.4656497	-.4656497

Los efectos de las variables en los distintos tipos de establecimientos son distintos

	Super Pequeños	Hiper	Super Grandes	Super Medianos
<i>retardo1</i>	0,4281685	0,6559388	0,6322404	0,5954268
<i>retardo4</i>	0,340386	0,0822364	0,1456201	0,2150342
<i>x9</i>	0,0092369	0,0252466	0,0163368	0,0112325
<i>x13</i>	0,0090225	0,003604	0,0065172	0,007365
<i>x17</i>	-0,0017983	0,0086818	0,0032448	0,0049844
<i>constante</i>	-0,4656497	-0,8346187	-0,6549531	-0,4570782

La variable x16 no aparece en el modelo porque ha sido eliminada por un problema de colinealidad, por lo tanto su efecto se recoge de forma conjunta en la estimación asociada a ella, no produciéndose diferencias por tipo de establecimiento.

Gráficamente los resultados obtenidos son:



El último paso es comprobar si la forma funcional es estable o bien es posible introducir alguna modificación que mejore la ecuación de regresión. Para comprobarlo calculamos nuevamente el contraste Reset.

Con la ecuación de regresión obtenida en primera estancia los resultados del contraste Reset son:

```
. : xtreg logproporción retardo1logproporción retardo4logproporción x9 x13 x17
> 6 x17 prediccion2, fe vce (cluster panel)
```

Fixed-effects (within) regression  
Group variable: panel

R-sq: within = 0.8025  
between = 0.9986  
overall = 0.9235

Number of obs = 416  
Number of groups = 4  
obs per group: min = 104  
avg = 104.0  
max = 104

corr(u\_i, Xb) = 0.7713

F(3,3) = .  
Prob > F = .

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.3773337	.1945901	1.94	0.148	-.2419388	.9966063
retardo4lo~n	.1619414	.0195467	8.28	0.004	-.0997351	.2241477
x9	.0091717	.0052632	1.74	0.180	-.0075782	.0259216
x13	.0059973	.0005243	11.44	0.001	.0043287	.0076659
x16	.00075	.0016722	0.45	0.684	-.0045717	.0060718
x17	.0023881	.0010763	2.22	0.113	-.001037	.0058132
prediccion2	-.072523	.0647055	-1.12	0.344	-.2784447	.1333986
_cons	-.7172685	.0977825	-7.34	0.005	-1.028456	-.4060809
sigma_u	.12312376					
sigma_e	.17130999					
rho	.34061173	(fraction of variance due to u_i)				

```
. testparm prediccion2
```

```
( 1) prediccion2 = 0
```

F( 1, 3) = 1.26  
Prob > F = 0.3440

A primera vista los resultados son óptimos, ya que el contraste nos indica que la variable generada, prediccion2 (la variable estimada al cuadrado), no es significativa. Pero a partir de este punto aparecen los problemas. Introducimos como variable explicativa prediccion3 (la variable estimada elevada al cubo) y mantenemos la variable prediccion2 y calculamos nuevamente la regresión; los resultados obtenidos son:

```
. . xtreg logproporcion retardo1logproporcion retardo4logproporcionÑ x9 x13 x1
> 6 x17 prediccion2 prediccion3, fe vce (cluster panel)
```

Fixed-effects (within) regression  
Group variable: panel

Number of obs = 416  
Number of groups = 4

R-sq: within = 0.8119  
between = 0.9998  
overall = 0.9256

Obs per group: min = 104  
avg = 104.0  
max = 104

corr(u\_i, xb) = 0.7768

F(3,3) = .  
Prob > F = .

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
retardo1lo~n	-.6614044	.4148649	-1.59	0.209	-1.98169 .6588808
retardo4lo~Ñ	-.3004775	.1113782	-2.70	0.074	-.6549326 .0539776
x9	-.0176966	.0098657	-1.79	0.171	-.0490937 .0137006
x13	-.0075463	.0030828	-2.45	0.092	-.0173572 .0022646
x16	-.0023852	.0008524	-2.80	0.068	-.0050979 .0003275
x17	-.0039126	.0024152	-1.62	0.204	-.0115989 .0037738
prediccion2	-1.076973	.3395372	-3.17	0.050	-2.157532 .0035863
prediccion3	-.1658424	.0549228	-3.02	0.057	-.3406312 .0089463
_cons	-.7556863	.0630633	-11.98	0.001	-.956382 -.5549906
sigma_u	.13023797				
sigma_e	.1673602				
rho	.37717176	(fraction of variance due to u_i)			

```
. testparm prediccion2 prediccion3
```

( 1) prediccion2 = 0  
( 2) prediccion3 = 0

F( 2, 3) = 7.54  
Prob > F = 0.0676

El contraste paramétrico nos informa de que no podemos rechazar la hipótesis de que ambas variables son 0; por lo tanto, existe en el modelo alguna variable o alguna no linealidad que no se ha tenido en cuenta. Lógicamente hay que buscarla.

Los primeros candidatos son las variables originales elevadas al cubo, pues recordemos que es en la segunda fase del contraste, cuando se ha incluido la variable prediccion3, cuando aparece el problema.

Se reformula el modelo introduciéndose las variables originales al cubo, y resultan algunas de ellas significativas, pero pueden eliminarse las variables **x9cubo** y

**x17cubo**, tal y como muestra el contraste, ya que no resultan estadísticamente significativas.

```
. . xtreg logproporcion retardo1logproporcion retardo4logproporcionN x9 x13 x1
> 6 x17 r1cubo r4cubo x9cubo x13cubo x16cubo x17cubo, fe vce (cluster panel)

Fixed-effects (within) regression              Number of obs   =       416
Group variable: panel                        Number of groups  =         4

R-sq:  within = 0.8085                      obs per group: min =       104
        between = 0.9992                      avg           =     104.0
        overall = 0.9272                      max           =       104

corr(u_i, xb) = 0.7800                      F(3,3)           =
                                                Prob > F         =
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.3485419	.071691	4.86	0.017	.1203892	.5766946
retardo4lo~N	.3508505	.0203929	17.20	0.000	.2859511	.4157499
x9	.0147973	.0017815	8.31	0.004	.0091277	.0204669
x13	.0050848	.0012215	4.16	0.025	.0011974	.0089722
x16	-.0062991	.0019044	-3.31	0.045	-.0123596	-.0002386
x17	.0041634	.0017606	2.36	0.099	-.0014395	.0097664
r1cubo	.0146323	.0044506	3.29	0.046	.0004686	.0287961
r4cubo	-.008748	.0017614	-4.97	0.016	-.0143537	-.0031423
x9cubo	-1.58e-06	1.50e-06	-1.06	0.368	-6.35e-06	3.18e-06
x13cubo	1.02e-06	3.72e-07	2.74	0.071	-1.64e-07	2.21e-06
x16cubo	.0000174	2.29e-06	7.61	0.005	.0000102	.0000247
x17cubo	-5.98e-07	4.24e-07	-1.41	0.253	-1.95e-06	7.51e-07
_cons	-.6727916	.0692219	-9.72	0.002	-.8930864	-.4524967
sigma_u	.1145481					
sigma_e	.16972283					
rho	.31295427	(fraction of variance due to u_i)				

```
. testparm x9cubo x17cubo
```

```
( 1) x9cubo = 0
( 2) x17cubo = 0
```

```
F( 2, 3) = 1.00
Prob > F = 0.4658
```

El paso siguiente es volver a comprobar si esta forma funcional ya es estable, aplicando nuevamente el contraste Reset pero ahora sobre esta nueva forma funcional.

```
. . xtreg logproporcion retardo1logproporcion retardo4logproporcionN x9 x13 x1
> 6 x17 r1cubo r4cubo x13cubo x16cubo prediccion2fe, fe vce (cluster panel)
```

```
Fixed-effects (within) regression      Number of obs   =      416
Group variable: panel                  Number of groups  =       4

R-sq:  within = 0.8119                  Obs per group:   min =      104
      between = 0.9992                      avg   =     104.0
      overall  = 0.9287                      max   =      104
```

```
corr(u_i, xb) = 0.7800                  F(3,3)          =
                                          Prob > F         =
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.2673142	.1469637	1.82	0.167	-.2003899	.7350183
retardo4lo~N	.0649423	.19156	0.34	0.757	-.5446873	.6745718
x9	-.0000738	.0110521	-0.01	0.995	-.0352467	.0350991
x13	-.0015509	.0051024	-0.30	0.781	-.017789	.0146872
x16	.0027315	.0056913	0.48	0.664	-.0153807	.0208438
x17	-.0004504	.0029728	-0.15	0.889	-.0099112	.0090103
r1cubo	-.024396	.0249496	-0.98	0.400	-.1037969	.0550049
r4cubo	-.0052958	.0024488	-2.16	0.119	-.0130888	.0024973
x13cubo	9.92e-07	4.93e-07	2.01	0.138	-5.78e-07	2.56e-06
x16cubo	-6.03e-06	.0000162	-0.37	0.735	-.0000577	.0000456
predicci~2fe	-.2575619	.1933984	-1.33	0.275	-.8730419	.357918
_cons	-.5377058	.0311049	-17.29	0.000	-.6366956	-.438716
sigma_u	.11250417					
sigma_e	.16798668					
rho	.30964299	(fraction of variance due to u_i)				

```
. testparm prediccion2fe
```

```
( 1) prediccion2fe = 0
```

```
      F( 1, 3) = 1.77
      Prob > F = 0.2751
```

Aparentemente los resultados son prometedores, pero no cantemos victoria hasta incluir la variable estimada al cubo.

Como se muestra en la tabla siguiente nuevamente el contraste Reset pone de manifiesto que algo falta en la forma funcional. Y este algo que falta ya no puede estar en las variables originales elevadas al cubo, pues ya están en el modelo; por lo tanto debemos buscarlo en las variables originales elevadas al cuadrado y si no es ahí, en las interacciones que se produzcan entre ellas.

```
. . xtreg logproporcion retardo1logproporcion retardo4logproporcion x9 x13 x1
> 6 x17 r1cubo r4cubo x13cubo x16cubo prediccion2fe prediccion3fe, fe vce (cl
> uster panel)
```

```
Fixed-effects (within) regression              Number of obs   =       416
Group variable: panel                          Number of groups =         4

R-sq:  within = 0.8205                        Obs per group:  min =       104
          between = 0.9998                      avg =      104.0
          overall = 0.9322                      max =       104

corr(u_i, Xb) = 0.7869                        F(3,3)          =
                                                Prob > F        =
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	-.4719917	.2223708	-2.12	0.124	-1.179675	.2356913
retardo4lo~n	-.5285383	.2210334	-2.39	0.097	-1.231965	.1748886
x9	-.0210508	.0093083	-2.26	0.109	-.0506738	.0085722
x13	-.0079026	.0036992	-2.14	0.122	-.019675	.0038698
x16	.012047	.0054972	2.19	0.116	-.0054476	.0295415
x17	-.0039051	.0022911	-1.70	0.187	-.0111965	.0033862
r1cubo	-.0223689	.0127157	-1.76	0.177	-.0628359	.018098
r4cubo	.0129726	.0061121	2.12	0.124	-.0064788	.0324239
x13cubo	-5.56e-07	4.29e-07	-1.30	0.285	-1.92e-06	8.08e-07
x16cubo	-.0000277	.0000126	-2.21	0.114	-.0000677	.0000122
predicci~2fe	-1.176947	.3145027	-3.74	0.033	-2.177835	-.176059
predicci~3fe	-.1806116	.0485619	-3.72	0.034	-.3351573	-.0260658
_cons	-.628751	.0775505	-8.11	0.004	-.8755513	-.3819506
sigma_u	.10981253					
sigma_e	.16431814					
rho	.30873081	(fraction of variance due to u_i)				

```
. testparm prediccion2fe prediccion3fe
```

```
( 1) prediccion2fe = 0
( 2) prediccion3fe = 0
```

```
F( 2, 3) = 7.00
Prob > F = 0.0741
```

Los resultados finales de esta búsqueda de la forma funcional más adecuada se muestran a continuación. En el modelo resultante se observa que se han eliminado algunas de las variables originales del modelo, así como algunas de las variables elevadas al cubo que en la especificación anterior aparecían, para ser sustituidas por otras nuevas.

```
. . xtreg logproporcion retardo1logproporcion retardo4logproporcion x9 x13 x1
> 6 x17 r1cubo r4cubo x13cubo x16cubo r1cuadrado r4cuadrado x9cuadrado x13cua
> drado x16cuadrado x17cuadrado, fe vce (cluster panel)
```

```
Fixed-effects (within) regression      Number of obs   =      416
Group variable: panel                  Number of groups  =       4
```

```
R-sq:  within = 0.8198      Obs per group: min =      104
       between = 0.9999      avg      =     104.0
       overall = 0.9319      max      =      104
```

```
corr(u_i, Xb) = 0.7875      F(3,3) = .
                          Prob > F = .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	-1.032388	.5322405	-1.94	0.148	-2.726215	.6614385
retardo4lo~n	-.2161548	.2421766	-0.89	0.438	-.9868689	.5545593
x9	.0176947	.0017358	10.19	0.002	.0121705	.0232189
x13	.0043465	.0020787	2.09	0.128	-.0022688	.0109618
x16	.0100538	.0045411	2.21	0.114	-.0043979	.0245056
x17	.0047197	.0018225	2.59	0.081	-.0010802	.0105197
r1cubo	-.0948026	.0477814	-1.98	0.141	-.2468642	.057259
r4cubo	-.0493504	.0207454	-2.38	0.098	-.1153715	.0166708
x13cubo	9.36e-07	1.17e-06	0.80	0.483	-2.80e-06	4.67e-06
x16cubo	.0000079	5.08e-06	15.55	0.001	.0000628	.0000952
r1cuadrado	-.6999177	.2866258	-2.44	0.092	-1.612089	.2122535
r4cuadrado	-.267334	.1300758	-2.06	0.132	-.6812931	.1466252
x9cuadrado	-.0001873	.0001435	-1.31	0.283	-.0006441	.0002694
x13cuadrado	.0000305	.0000935	0.33	0.766	-.000267	.0003279
x16cuadrado	-.0020923	.0002345	-8.92	0.003	-.0028385	-.0013462
x17cuadrado	-.0000373	.000032	-1.17	0.327	-.0001391	.0000644
_cons	-1.878147	.4248621	-4.42	0.021	-3.230248	-.5260464
sigma_u	.11024259					
sigma_e	.16548057					
rho	.30739187	(fraction of variance due to u_i)				

```
. testparm x13cuadrado x17cuadrado x9cuadrado r4cuadrado r1cubo x16 retardo1l
> ogproporcion
```

```
( 1) retardo1logproporcion = 0
( 2) x16 = 0
( 3) r1cubo = 0
( 4) r4cuadrado = 0
( 5) x9cuadrado = 0
( 6) x13cuadrado = 0
( 7) x17cuadrado = 0
Constraint 2 dropped
Constraint 5 dropped
Constraint 6 dropped
Constraint 7 dropped
```

```
F( 3, 3) = 2.27
Prob > F = 0.2592
```

Una vez eliminadas las variables no significativas aplicamos el contraste Reset sobre la ecuación resultante.

Nuevamente, cuando incluimos en el modelo la variable estimada al cuadrado, resulta no ser significativa; cosa que ya se había producido en los casos anteriores, el problema será cuando se incluya la variable al cubo.



```
. . xtreg logproporcion retardo4logproporcionN x9 x13 x17 r4cubo x13cubo x16cu
> bo r1cuadrado x16cuadrado prediccionfe3cuadrado, fe vce (cluster panel)
```

```
Fixed-effects (within) regression      Number of obs   =      416
Group variable: panel                  Number of groups =       4
```

```
R-sq:  within = 0.8105      Obs per group: min =      104
        between = 0.9991      avg      =     104.0
        overall = 0.9283      max      =      104
```

```
corr(u_i, Xb) = 0.7792      F(3,3) = .
                          Prob > F = .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo4lo~N	.401195	.0734958	5.46	0.012	.1672986	.6350913
x9	.0158481	.0044008	3.60	0.037	.0018426	.0298535
x13	.0058548	.0023832	2.46	0.091	-.0017296	.0134392
x17	.0031411	.0015511	2.03	0.136	-.0017952	.0080775
r4cubo	-.0105359	.0022661	-4.65	0.019	-.0177477	-.0033241
x13cubo	1.15e-06	3.43e-07	3.36	0.044	6.22e-08	2.25e-06
x16cubo	.0000644	.0000107	6.01	0.009	.0000303	.0000985
r1cuadrado	-.1500238	.0347338	-4.32	0.023	-.2605623	-.0394852
x16cuadrado	-.0013786	.0002155	-6.40	0.008	-.0020644	-.0006928
prediccion~do	.0296749	.060742	0.49	0.659	-.1636332	.2229831
_cons	-.9214565	.0316838	-29.08	0.000	-1.022289	-.8206245
sigma_u	.1113344					
sigma_e	.16841842					
rho	.30410488	(fraction of variance due to u_i)				

```
. testparm prediccionfe3cuadrado
```

```
( 1) prediccionfe3cuadrado = 0
```

```
F( 1, 3) = 0.24
Prob > F = 0.6587
```

El resultado de incluir al variable estimada al cubo se muestra en el cuadro siguiente.

Como puede comprobarse, ahora sí podemos aceptar la hipótesis nula respecto a la forma funcional. Por lo tanto esta forma ya es estable y no existe ningún problema de no linealidad.

La interpretación de las variables continúa siendo la original, el problema es calcular el efecto de estas variables sobre la variable dependiente, que ha dejado de ser constante, ya que cuando queremos calcular el efecto marginal dependerá de los valores de las distintas variables.

```
. . xtreg logproporción retardo4logproporción x9 x13 x17 r4cubo x13cubo x16cu
> bo r1cuadrado x16cuadrado prediccionfe3cuadrado prediccionfe3cubo, fe vce
> (cluster panel)
```

```
Fixed-effects (within) regression      Number of obs   =      416
Group variable: panel                  Number of groups =        4

R-sq:  within = 0.8191                  Obs per group:  min =      104
        between = 0.9998                  avg =     104.0
        overall = 0.9317                  max =      104

corr(u_i, Xb) = 0.7868                  F(3,3) = .
                                         Prob > F = .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo4lo~ñ	-.4462528	.3038977	-1.47	0.238	-1.413391	.5208854
x9	-.0149598	.0115098	-1.30	0.285	-.051589	.0216695
x13	-.0052085	.0044364	-1.17	0.325	-.0193271	.0089101
x17	-.0023996	.002627	-0.91	0.428	-.01076	.0059608
r4cubo	.0141822	.0094303	1.50	0.230	-.0158294	.0441937
x13cubo	-6.88e-07	6.27e-07	-1.10	0.353	-2.68e-06	1.31e-06
x16cubo	-.0000735	.0000501	-1.47	0.239	-.0002329	.000086
r1cuadrado	.1255987	.1013963	1.24	0.304	-.1970895	.448287
x16cuadrado	.0017112	.0011252	1.52	0.226	-.0018698	.0052922
predicció~do	-1.022989	.3922126	-2.61	0.080	-2.271184	.2252066
predicció~bo	-.1661408	.0626493	-2.65	0.077	-.3655188	.0332373
_cons	-.4033123	.2280127	-1.77	0.175	-1.12895	.322326
sigma_u	.10953089					
sigma_e	.16474332					
rho	.30653601	(fraction of variance due to u_i)				

```
. testparm prediccionfe3cuadrado prediccionfe3cubo
```

```
( 1) prediccionfe3cuadrado = 0
( 2) prediccionfe3cubo = 0
```

```
F( 2, 3) = 3.65
Prob > F = 0.1572
```

Así pues, el modelo final está formado por las variables retardo 4, x9, x13, x17, el retardo 4 elevado al cubo, x13 al cubo, x16 al cubo, retardo 1 al cuadrado y x16 al cuadrado. Esta forma funcional no presenta ningún problema de variables omitidas, tal y como nos indica el contraste Reset.

Por tanto y eliminando las variables introducidas en el contraste, tenemos la estructura definitiva estimada para el modelo.

```
. xtreg logproporc~n retardo4logproporc~n x9 x13 x17 r4cubo x13cubo x16cubo r1cuadrado x
> 16cuadrado, fe vce (cluster panel)
```

```
Fixed-effects (within) regression      Number of obs      =      416
Group variable: panel                  Number of groups    =         4

R-sq:  within  = 0.8103                  Obs per group: min =      104
      between  = 0.9989                      avg      =    104.0
      overall  = 0.9283                      max      =      104
```

```
corr(u_i, Xb)  = 0.7776                  F(3,3)              =      .
                                          Prob > F             =      .
```

(Std. Err. adjusted for 4 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo4lo~n	.3815741	.0391013	9.76	0.002	.2571363	.5060118
x9	.0142113	.0013358	10.64	0.002	.0099602	.0184624
x13	.0050983	.0012955	3.94	0.029	.0009756	.0092211
x17	.0027367	.0007862	3.48	0.040	.0002346	.0052387
r4cubo	-.010813	.0026524	-4.08	0.027	-.0192541	-.0023718
x13cubo	1.17e-06	3.43e-07	3.43	0.042	8.35e-08	2.27e-06
x16cubo	.0000582	.0000101	5.78	0.010	.0000262	.0000903
r1cuadrado	-.1312634	.0073999	-17.74	0.000	-.154813	-.1077138
x16cuadrado	-.0012601	.0002633	-4.79	0.017	-.002098	-.0004222
_cons	-.9038911	.0145269	-62.22	0.000	-.9501223	-.8576598
sigma_u	.11083157					
sigma_e	.16830019					
rho	.30248861	(fraction of variance due to u_i)				

## 6.6.-ESTIMACIÓN DEL MODELO DINÁMICO DE EFECTOS INDIVIDUALES ESPECÍFICOS TIPO DE ESTABLECIMIENTO HIPERMERCADOS.

En este modelo, que es el único por tipo de establecimiento, procederemos de forma similar a los anteriores. Es de destacar que es el único modelo donde la agrupación se produce por tipo de establecimiento; mientras que en los anteriores la agrupación se ha producido por formato. Como decíamos en el capítulo anterior, esta agrupación hace referencia a la fuerza que tienen los hipermercados en el mercado español que supera las políticas de los propios fabricantes, imponiendo un tratamiento comercial específico para este tipo de establecimientos.

Dicho esto, vamos a determinar que variables son las más influyentes en este modelo.

```
. xtreg logproporción retardo1logproporción retardo2logproporción retardo3logproporción re
> tardo4logproporción x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 x13 x14 x15 x16 x17 x18 x19 x20
> x21 x22 x23 x24 x25 x26 x27 x28 x29 x30 x31 x32 x33 x34 x35, fe vce(cluster panel)
```

```
Fixed-effects (within) regression              Number of obs   =       624
Group variable: panel                        Number of groups =         6

R-sq:  within = 0.4435                      Obs per group: min =      104
       between = 0.9925                      avg =      104.0
       overall = 0.9683                      max =      104

                                F(5,5) = .
corr(u_i, xb) = 0.9467                    Prob > F = .
```

(Std. Err. adjusted for 6 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.4624004	.0867936	5.33	0.003	.2392905	.6855104
retardo2lo~n	.0227866	.0721277	0.32	0.765	-.1626237	.2081968
retardo3lo~n	.0359647	.0101745	3.53	0.017	.0098104	.0621191
retardo4lo~n	.1460976	.0584512	2.50	0.055	-.004156	.2963513
x1	-.0020106	.0008445	-2.38	0.063	-.0041814	.0001601
x2	.3267293	.4578003	0.71	0.507	-.850084	1.503543
x3	-.013568	.2127943	-0.06	0.952	-.560573	.533437
x4	-.3577372	1.027938	-0.35	0.742	-3.000135	2.284661
x5	.0000846	.001024	0.08	0.937	-.0025476	.0027169
x6	-.0011557	.0014293	-0.81	0.456	-.00483	.0025185
x7	.0002226	.0013309	0.17	0.874	-.0031987	.0036439
x8	.3555681	.217105	1.64	0.162	-.202518	.9136542
x9	.0044579	.002852	1.56	0.179	-.0028733	.0117892
x10	.0076802	.001746	4.40	0.007	.0031919	.0121686
x11	.0029037	.001445	2.01	0.101	-.0008107	.0066181
x12	-.0014206	.0011805	-1.20	0.283	-.0044552	.001614
x13	.0029177	.0022817	1.28	0.257	-.0029475	.008783
x14	-.0012846	.0007555	-1.70	0.150	-.0032265	.0006574
x15	-.5172778	.6234187	-0.83	0.444	-2.119827	1.085271
x16	-.0021089	.002018	-1.05	0.344	-.0072963	.0030785
x17	-.0006934	.0008762	-0.79	0.465	-.0029456	.0015589
x18	-.000495	.0008045	-0.62	0.565	-.0025631	.001573
x19	-.0000583	.0015634	-0.04	0.972	-.0040772	.0039606
x20	.0017102	.0011866	1.44	0.209	-.0013401	.0047605
x21	.0007603	.0004816	1.58	0.175	-.0004776	.0019982
x22	.2140629	.1112432	1.92	0.112	-.0718968	.5000226
x23	-.0009607	.0011657	-0.82	0.447	-.0039573	.0020359
x24	-.0031353	.0012792	-2.45	0.058	-.0064236	.000153
x25	.001283	.0004728	2.71	0.042	.0000677	.0024983
x26	-.0013533	.0005119	-2.64	0.046	-.0026693	-.0000373
x27	-.0020986	.0009366	-2.24	0.075	-.0045063	.000309
x28	.0019944	.0044225	0.45	0.671	-.0093739	.0133626
x29	-.1409511	.0768985	-1.83	0.126	-.3386251	.0567228
x30	.0080292	.0036548	2.20	0.079	-.0013657	.0174241
x31	-.0037134	.0059405	-0.63	0.559	-.0189841	.0115572
x32	-.0015538	.0006536	-2.38	0.063	-.0032339	.0001263
x33	-.0010106	.0017898	-0.56	0.597	-.0056113	.0035901
x34	-.0005107	.000548	-0.93	0.394	-.0019193	.000898
x35	-.054547	.0203237	-2.68	0.044	-.1067908	-.0023032
_cons	-.7959293	.742054	-1.07	0.332	-2.70344	1.111581
sigma_u	.49211793					
sigma_e	.21016367					
rho	.84575188	(fraction of variance due to u_i)				

La primera aproximación al modelo nos indica un elevado número de variables que no son significativas, por tanto debemos ir purgando el modelo hasta quedarnos con aquellas variables que sean significativas tanto desde el punto de vista estadístico y como desde el punto de vista económico.

Para eliminar las variables no significativas hacemos un test de significación conjunto.

```
. . testparm x1 x2 x3 x4 x5 x6 x7 x12 x13 x14 x15 x16 x17 x18 x19 x20 x21 x23 x
> 25 x28 x31 x34

( 1) x1 = 0
( 2) x2 = 0
( 3) x3 = 0
( 4) x4 = 0
( 5) x5 = 0
( 6) x6 = 0
( 7) x7 = 0
( 8) x12 = 0
( 9) x13 = 0
(10) x14 = 0
(11) x15 = 0
(12) x16 = 0
(13) x17 = 0
(14) x18 = 0
(15) x19 = 0
(16) x20 = 0
(17) x21 = 0
(18) x23 = 0
(19) x25 = 0
(20) x28 = 0
(21) x31 = 0
(22) x34 = 0
Constraint 1 dropped
Constraint 5 dropped
Constraint 6 dropped
Constraint 7 dropped
Constraint 8 dropped
Constraint 9 dropped
Constraint 10 dropped
Constraint 12 dropped
Constraint 13 dropped
Constraint 14 dropped
Constraint 15 dropped
Constraint 16 dropped
Constraint 17 dropped
Constraint 18 dropped
Constraint 19 dropped
Constraint 20 dropped
Constraint 22 dropped

F( 5, 5) = 1.91
Prob > F = 0.2471
```

Como puede apreciarse, el p valor asociado al contraste, nos indica que la hipótesis nula no se rechaza y por lo tanto, es posible eliminar las variables del modelo.

```
. . xtreg logproporcion retardo1logproporcion retardo2logproporcionN retardo3logproporcionN retardo4logproporcionN x8 x9 x10 x11 x22 x24 x26 x27 x29 x30 x31 x32 x33 x35, fe vce(cluster panel)
```

```
Fixed-effects (within) regression               Number of obs   =       624
Group variable: panel                          Number of groups =        6

R-sq:  within = 0.4248                          Obs per group: min =       104
        between = 0.9983                          avg =      104.0
        overall = 0.9739                          max =       104

corr(u_i, Xb) = 0.9740                          F(5,5) =
                                                Prob > F =
```

(Std. Err. adjusted for 6 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.4726588	.1044343	4.53	0.006	.2042019	.7411156
retardo2lo~N	.0179962	.0760602	0.24	0.822	-.1775227	.2135151
retardo3lo~N	.0312578	.0097994	3.19	0.024	.0060677	.0564478
retardo4lo~N	.1367346	.0523721	2.61	0.048	.0021079	.2713613
x8	.2706562	.0763797	3.54	0.016	.074316	.4669965
x9	.0064558	.0026765	2.41	0.061	-.0004245	.013336
x10	.0076103	.0006893	11.04	0.000	.0058385	.0093822
x11	.0046845	.0011215	4.18	0.009	.0018015	.0075676
x22	.2486798	.1138453	2.18	0.081	-.0439689	.5413284
x24	-.0029014	.0011255	-2.58	0.050	-.0057945	-.8.27e-06
x26	-.0012586	.000159	-7.92	0.001	-.0016674	-.0008499
x27	-.0015814	.0007384	-2.14	0.085	-.0034795	.0003168
x29	-.1320331	.0850104	-1.55	0.181	-.3505592	.0864931
x30	.0061217	.0024184	2.53	0.052	-.000095	.0123384
x32	-.0015037	.0001527	-9.85	0.000	-.0018962	-.0011113
x33	-.0023178	.0007063	-3.28	0.022	-.0041335	-.0005022
x35	-.0583936	.0252972	-2.31	0.069	-.1234221	.006635
_cons	-1.367683	.0942231	-14.52	0.000	-1.609891	-1.125475
sigma_u	.44986869					
sigma_e	.20970511					
rho	.82149465	(fraction of variance due to u_i)				

En este modelo aparecen como no significativas individualmente, variables como retardo 2 y x29; son variables que si se incluyen en el contraste anterior rechazaríamos la hipótesis nula. La explicación a este hecho hay que buscarla en que estas variables, no significativas individualmente, aumentan notablemente la suma de cuadrados de residuos del modelo restringido llevando a rechazar la hipótesis. La razón puede ser un problema de multicolinealidad entre las variables. Además, y reafirmando esta idea, cuando sometemos al modelo nuevo al mismo tipo de contraste, los resultados nos indican que estas variables pueden eliminarse del modelo, ahora bien, este modelo es distinto al primero que habíamos estado utilizando.

```
. . testparm retardo2logproporcionN x29
```

```
( 1) retardo2logproporcionN = 0
( 2) x29 = 0
```

```
F( 2, 5) = 1.32
Prob > F = 0.3461
```

```
. xtreg logproporcion retardo1logproporcion retardo3logproporcion retardo4logp
> roporcion x8 x9 x10 x11 x22 x24 x26 x27 x30 x32 x33 x35, fe vce(cluster panel
> )
```

```
Fixed-effects (within) regression              Number of obs   =       624
Group variable: panel                          Number of groups  =         6

R-sq:  within = 0.4238                        Obs per group: min =       104
        between = 0.9981                      avg           =     104.0
        overall = 0.9735                      max           =       104

                                E(5,5)         =         .
corr(u_i, xb) = 0.9730                    Prob > F         =         .
```

(Std. Err. adjusted for 6 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.4803888	.0747718	6.42	0.001	.2881818	.6725957
retardo3lo~n	.0383177	.0366544	1.05	0.344	-.0559056	.1325409
retardo4lo~n	.1360769	.0557631	2.44	0.059	-.0072667	.2794206
x8	.282788	.059055	4.79	0.005	.1309823	.4345937
x9	.0064425	.002589	2.49	0.055	-.0002128	.0130977
x10	.0075847	.0006748	11.24	0.000	.00585	.0093193
x11	.0047225	.0011297	4.18	0.009	.0018184	.0076265
x22	.2498862	.1101947	2.27	0.073	-.0333783	.5331507
x24	-.0028631	.0011457	-2.50	0.055	-.0058082	.0000821
x26	-.001289	.0001764	-7.31	0.001	-.0017424	-.0008355
x27	-.0016363	.000742	-2.21	0.079	-.0035437	.0002711
x30	.0061191	.0023378	2.62	0.047	.0001095	.0121286
x32	-.0015804	.0001648	-9.59	0.000	-.002004	-.0011567
x33	-.0022047	.0005439	-4.05	0.010	-.0036029	-.0008065
x35	-.0602965	.0251486	-2.40	0.062	-.1249431	.0043501
_cons	-1.517354	.0732534	-20.71	0.000	-1.705658	-1.329051
sigma_u	.45757775					
sigma_e	.20954018					
rho	.82664912	(fraction of variance due to u_i)				

El nuevo modelo presenta aún otro tipo de problemas. En este caso estamos hablando de variables que son significativas estadísticamente, pero económicamente presentan problemas.

Este es el caso de las variables x8, x22 y x30, que tienen un signo asociado distinto al esperado. Utilizando un principio de prudencia económica, se prefiere eliminar las variables del modelo aunque perdamos la explicación que aportan sobre la variable objeto de estudio. Además, variables que en el modelo anterior, eran significativas individualmente, dejan de serlo como consecuencia de la multicolinealidad, este es el caso de la variable retardo 3.

```

. . testparm retardo3logproporcionN
( 1) retardo3logproporcionN = 0
      F( 1,      5) =      1.09
      Prob > F =      0.3437

. . testparm retardo3logproporcionN x8 x22 x30
( 1) retardo3logproporcionN = 0
( 2) x8 = 0
( 3) x22 = 0
( 4) x30 = 0
      F( 4,      5) =     18.23
      Prob > F =      0.0035

```

Puede apreciarse que la variable retardo 3 es posible eliminarla del modelo, pero no así con las variables x8, 22 y 30, que las eliminamos con una pérdida del poder explicativo del modelo.

```

Fixed-effects (within) regression               Number of obs   =      624
Group variable: panel                          Number of groups =       6

R-sq:  within = 0.4136                        obs per group: min =     104
        between = 0.9983                      avg           =    104.0
        overall = 0.9730                      max           =     104

corr(u_i, Xb) = 0.9749                        F(5,5)          =      .
                                                Prob > F         =      .

```

(Std. Err. adjusted for 6 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.4963589	.0724523	6.85	0.001	.3101143	.6826035
retardo4lo~N	.1550381	.068407	2.27	0.073	-.0208076	.3308838
x9	.0061658	.0026669	2.31	0.069	-.0006897	.0130214
x10	.0074972	.0007713	9.72	0.000	.0055147	.0094798
x11	.0047561	.0009197	5.17	0.004	.002392	.0071203
x24	-.0028202	.0011406	-2.47	0.056	-.0057523	.0001118
x26	-.0010821	.0001735	-6.24	0.002	-.001528	-.0006362
x27	-.0011519	.0004082	-2.82	0.037	-.0022013	-.0001025
x32	-.0011899	.0002677	-4.44	0.007	-.0018781	-.0005017
x33	-.001583	.0002585	-6.12	0.002	-.0022476	-.0009184
x35	-.0503941	.0269027	-1.87	0.120	-.1195496	.0187615
_cons	-1.034644	.0412353	-25.09	0.000	-1.140643	-.9286453
sigma_u	.48114194					
sigma_e	.21069178					
rho	.839098	(fraction of variance due to u_i)				

En este modelo, otra variable que hasta ahora era significativa, deja de serlo (nuevamente multicolinealidad) y tiene que ser eliminada.

```

. testparm x35
( 1) x35 = 0
      F( 1,      5) =      3.51
      Prob > F =      0.1199

```



La formulación más simple del modelo es:

```
. xtreg logproporción retardo1logproporción retardo4logproporción x9 x10 x11 x24 x26 x27 x3
> 2 x33, fe vce (cluster panel)
```

Fixed-effects (within) regression	Number of obs	=	624
Group variable: panel	Number of groups	=	6
R-sq: within = 0.4124	obs per group: min =		104
between = 0.9984	avg =		104.0
overall = 0.9731	max =		104
corr(u_i, xb) = 0.9754	F(5,5)	=	.
	Prob > F	=	.

(Std. Err. adjusted for 6 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
retardo1lo~n	.4981634	.073581	6.77	0.001	.3090175 .6873093
retardo4lo~n	.1569713	.0670535	2.34	0.066	-.0153953 .3293378
x9	.006346	.0025935	2.45	0.058	-.0003208 .0130127
x10	.0072708	.0008382	8.67	0.000	.0051161 .0094254
x11	.0044697	.0008105	5.52	0.003	.0023863 .0065531
x24	-.0027524	.0011512	-2.39	0.062	-.0057115 .0002068
x26	-.0010845	.000176	-6.16	0.002	-.0015369 -.0006321
x27	-.0011324	.0004035	-2.81	0.038	-.0021696 -.0000951
x32	-.001221	.0002502	-4.88	0.005	-.0018641 -.0005779
x33	-.0015896	.0002555	-6.22	0.002	-.0022464 -.0009328
_cons	-1.023057	.040973	-24.97	0.000	-1.128382 -.9177328
sigma_u	.47651382				
sigma_e	.21073385				
rho	.83641628	(fraction of variance due to u_i)			

Que está compuesto por las variables siguientes:

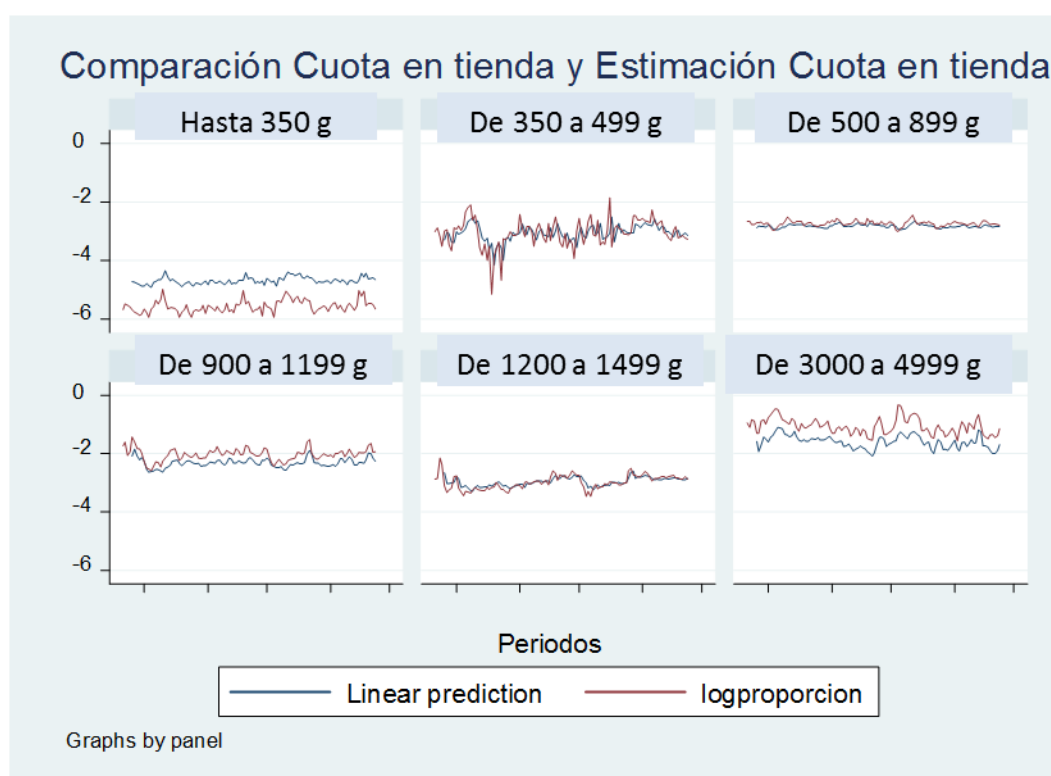
1. **Retardo1** Esta variable mide la influencia que tiene la variable cuota de mercado del periodo anterior. El signo asociado es positivo de tal forma que un aumento de la cuota de mercado en el periodo anterior produce un aumento en la variable a explicar del periodo actual.
2. **Retardo4** Mide la influencia de la variable cuota de mercado de 4 periodos atrás. El signo asociado es positivo de manera que la relación entre la variable a explicar y esta variable es directa; aumentos en una suponen aumentos en la otra y viceversa.
3. **Cobertura Promoción Folleto y Exposición Especial Marca 1: X<sub>9</sub>** Esta variable mide la influencia de este tipo de acción en la cuota en tienda. El signo esperado es positivo, como realmente se produce en la estimación; por lo que podemos afirmar que esta acción produce resultados positivos en las ventas de la Marca 1.

4. **Cobertura Promoción sólo Exposición Especial Marca 1:  $X_{10}$**  Esta variable recoge idéntica información a la variable anterior, pero referida exclusivamente a exposición especial. El signo asociado es positivo, como el esperado, con lo que nos indica que la utilización de esta acción produce un efecto positivo sobre la cuota en tienda de la Marca 1.
5. **Cobertura Promoción solo Folleto Marca 1:  $X_{11}$**  Esta variable recoge el efecto individual positivo de la promoción de Folleto sobre la cuota en tienda de la Marca 1.
6. **Cobertura Promoción solo Exposición Especial Resto de Fabricantes:  $X_{24}$**  Esta variable mide la relación existente entre un tipo de promoción realizada por la competencia agrupada bajo la enseña Resto de Fabricantes y la cuota en tienda de la Marca 1. El signo asociado a ella es NEGATIVO de forma que a mayor actividad promocional de la competencia menor será la cuota en tienda de la Marca 1.
7. **Cobertura Otras Promociones Resto de Fabricantes:  $X_{26}$**  Mide la relación entre otros tipos de promociones realizadas por el resto de los fabricantes y la cuota en tienda de la Marca 1. El signo esperado es NEGATIVO, ya que es esperable que una mayor actividad promocional del resto de fabricantes esté relacionada con una menor cuota en tienda de la Marca 1.
8. **Cobertura Reducción Temporal de Precios Resto de Fabricantes:  $X_{27}$**  Esta variable mide una relación muy similar a la anterior; la diferencia está en el tipo de promociones consideradas. El signo esperado es nuevamente NEGATIVO ya que se trata de una relación entre la Marca 1 y su competencia representada por el Resto de Fabricantes.
9. **Cobertura Promoción Solo Folleto Marcas Distribución:  $X_{32}$**  La variable que consideramos en esta ocasión mide la relación establecida entre la intensidad promocional, para el tipo “SÓLO FOLLETO”, realizada por las marcas de la distribución y la Marca 1. Es esperable que esta relación sea inversa, de tal modo que a mayor actividad promocional de las marcas de distribución, menor sea la

cuota en tienda de la Marca 1. Esta relación inversa se materializa en un beta con signo NEGATIVO, como puede apreciarse en la salida del modelo.

10. **Cobertura Otras Promociones Marcas Distribución:  $X_{33}$**  La relación existente en esta variable es simétrica a la anterior y por las mismas razones, su coeficiente estimado debería tener y tiene signo NEGATIVO.

A continuación, en el gráfico, vemos la comparación entre las cuotas de ventas estimadas y reales en los distintos formatos considerados en el canal de hipermercados.



En el gráfico se aprecia como los niveles de ajuste de la variable son distintos dependiendo del formato que estemos considerando; así los formatos hasta 350 gr. y De 3000 a 4999 gr. son los que mayores discrepancias presentan entre la variable real y la estimada.

Ahora comprobaremos si los efectos de las variables independientes sobre la variable a explicar son los mismos según los diferentes formatos. En esta prueba

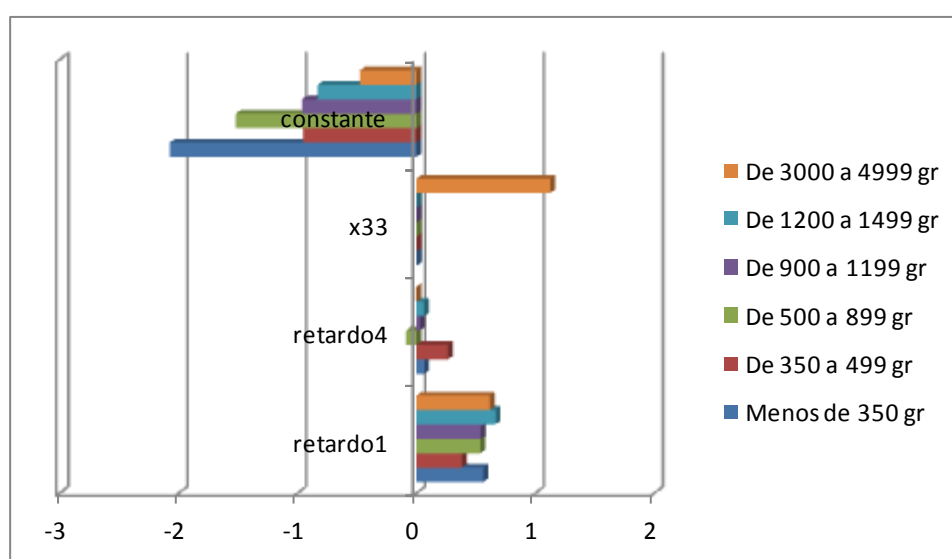
realizamos una regresión sobre la variable a explicar respecto a las variables que han resultado significativas en el modelo y les añadimos: la interacción entre las variables y los distintos formatos más las variables ficticias “formatos”.

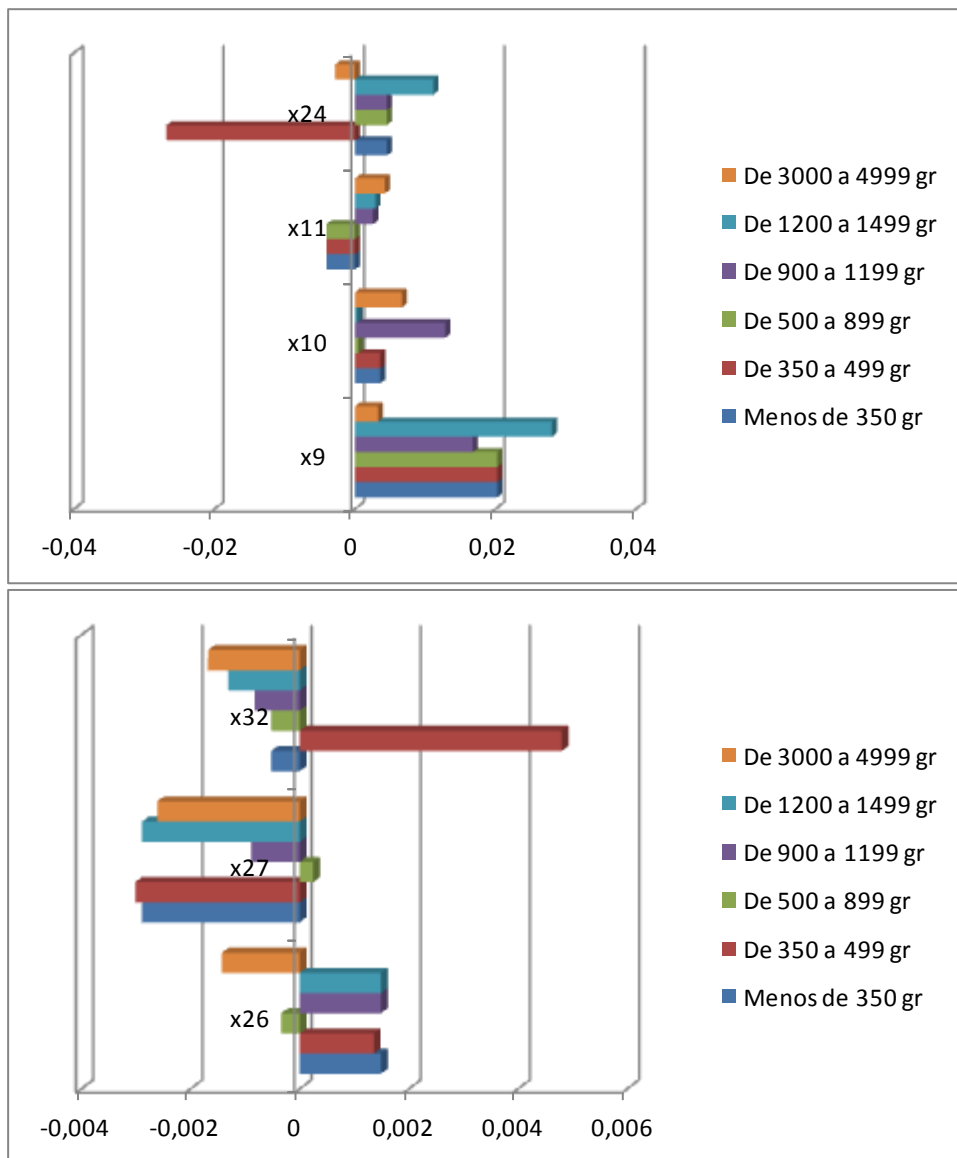
Los resultados nos muestran que efectivamente, los efectos de las variables son distintos según el formato considerado.

	Menos de 350 gr	De 350 a 499 gr	De 500 a 899 gr	De 900 a 1199 gr	De 1200 a 1499 gr	De 3000 a 4999 gr
<i>retardo1</i>	0,5600678	0,3790489	0,5365148	0,5420783	0,6642947	0,6161422
<i>retardo4</i>	0,0639652	0,2649013	-0,0883591	0,0288319	0,0629037	-0,0040463
<i>x9</i>	0,0202269	<b>0,0202269</b>	<b>0,0202269</b>	0,0167499	0,0280298	0,0032663
<i>x10</i>	0,003583	<b>0,003583</b>	0,0004587	0,0128433	0,0003095	0,0066794
<i>x11</i>	-0,004024	<b>-0,004024</b>	<b>-0,004024</b>	0,0025843	0,00289	0,0042635
<i>x24</i>	0,0045491	-0,0268494	<b>0,0045491</b>	<b>0,0045491</b>	0,0111359	-0,0028414
<i>x26</i>	0,0014906	0,0013589	-0,0003371	<b>0,0014906</b>	<b>0,0014906</b>	-0,0014245
<i>x27</i>	-0,0028865	-0,003012	0,0002485	-0,0008855	<b>-0,0028865</b>	-0,0026102
<i>x32</i>	-0,00052	0,0047993	<b>-0,00052</b>	-0,0008235	-0,0013078	-0,0016782
<i>x33</i>	0,0022876	<b>0,0022876</b>	<b>0,0022876</b>	<b>0,0022876</b>	<b>0,0022876</b>	1,1224956
<i>constante</i>	-2,076173	-0,955965	-1,5193152	-0,961323	-0,832898	-0,476014

En rojo están los efectos de las variables, eliminadas por problemas de colinealidad, y que en ese caso son iguales, ya que resulta imposible distinguir entre unos y otros.

Gráficamente, los resultados son:





Por último, se trata de comprobar si la forma funcional elegida es correcta, o sí por el contrario, existe alguna no linealidad que es necesario incluir en el modelo.

Con la ecuación de regresión obtenida en primera estancia los resultados del contraste Reset son:

```
. xtreg logproporcion retardo1logproporcion retardo4logproporcionÑ x9 x10 x11 x
> 24 x26 x27 x32 x33 prediccion2, fe vce(cluster panel)
```

```
Fixed-effects (within) regression      Number of obs   =      624
Group variable: panel                 Number of groups =       6
```

```
R-sq:  within = 0.4128      obs per group: min =      104
      between = 0.9964      avg =      104.0
      overall = 0.9710      max =      104
```

```
corr(u_i, Xb) = 0.9659      F(5,5) = .
                          Prob > F = .
```

(Std. Err. adjusted for 6 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.5617166	.0530337	10.59	0.000	.4253891	.6980442
retardo4lo~Ñ	.1764078	.0763285	2.31	0.069	-.0198008	.3726164
x9	.006588	.0025698	2.56	0.050	-.0000179	.0131939
x10	.0075024	.0008629	8.69	0.000	.0052841	.0097206
x11	.0049498	.0009572	5.17	0.004	.0024892	.0074104
x24	-.0028143	.001193	-2.36	0.065	-.0058809	.0002524
x26	-.0010569	.000151	-7.00	0.001	-.0014451	-.0006688
x27	-.0012739	.0004332	-2.94	0.032	-.0023874	-.0001603
x32	-.0013404	.0002255	-5.94	0.002	-.0019201	-.0007608
x33	-.0017291	.0003153	-5.48	0.003	-.0025397	-.0009186
prediccion2	.0209582	.0143295	1.46	0.203	-.0158769	.0577933
_cons	-.9794587	.0352377	-27.80	0.000	-1.07004	-.8888772
sigma_u	.48946122					
sigma_e	.21084304					
rho	.84348398	(fraction of variance due to u_i)				

```
. testparm prediccion2
```

```
( 1) prediccion2 = 0
```

```
      F( 1, 5) = 2.14
      Prob > F = 0.2034
```

En este primer paso la variable añadida, predicción2; que no es otra cosa que las estimaciones al cuadrado, resultan ser no significativas.

Queda un último paso, y es saber si existe influencia de alguna variable de orden 3 o de los productos cruzados de alguna de orden 2 con otras de orden 1. Para ello introducimos en el modelo las variables prediccion2 y prediccion3 (son las estimaciones del modelo al cubo).

Los resultados de la regresión se muestran a continuación, y puede observarse que ambas variables no son significativas individualmente. Aun así hacemos un contraste conjunto para comprobar esta hipótesis. Y en este contraste conjunto no rechazamos la hipótesis nula, esto es que los  $\beta$  asociados a ellas son 0; por tanto podemos concluir que la forma funcional es correcta y no hay ninguna no linealidad no incluida.

```
. xtreg logproporcion retardo1logproporcion retardo4logproporcionN x9 x10 x11 x
> 24 x26 x27 x32 x33 prediccion2 prediccion3, fe vce(cluster panel)
```

```
Fixed-effects (within) regression      Number of obs   =      624
Group variable: panel                 Number of groups  =       6

R-sq:  within = 0.4128                Obs per group: min =     104
      between = 0.9963                  avg   =    104.0
      overall  = 0.9709                  max   =     104

                                F(5,5)      =      .
corr(u_i, Xb) = 0.9655              Prob > F      =      .
```

(Std. Err. adjusted for 6 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.5134668	.1648301	3.12	0.026	.0897576	.937176
retardo4lo~n	.1610654	.0891947	1.81	0.131	-.0682169	.3903476
x9	.0061035	.0021428	2.85	0.036	.0005953	.0116116
x10	.0069691	.0016957	4.11	0.009	.0026102	.011328
x11	.0044916	.0018128	2.48	0.056	-.0001683	.0091515
x24	-.0026311	.001329	-1.98	0.105	-.0060474	.0007852
x26	-.0009947	.0002603	-3.82	0.012	-.0016637	-.0003256
x27	-.0011658	.0006232	-1.87	0.120	-.0027678	.0004362
x32	-.0012164	.0003696	-3.29	0.022	-.0021664	-.0002663
x33	-.0015522	.0007193	-2.16	0.083	-.0034012	.0002968
prediccion2	-.0135526	.1153709	-0.12	0.911	-.3101229	.2830176
prediccion3	-.0038017	.0127701	-0.30	0.778	-.0366284	.0290249
_cons	-.9649317	.0729179	-13.23	0.000	-1.152373	-.7774902
sigma_u	.49098719					
sigma_e	.2110095					
rho	.84409653	(fraction of variance due to u_i)				

```
. testparm prediccion2 prediccion3
```

```
( 1) prediccion2 = 0
( 2) prediccion3 = 0
```

```
      F( 2,      5) =      0.71
      Prob > F =      0.5332
```

Así pues el modelo estimado queda como sigue:

```
. xtreg logproporcion retardo1logproporcion retardo4logproporcion x9 x10 x11 x24 x26 x27 x3
> 2 x33, fe vce(cluster panel)
```

```
Fixed-effects (within) regression               Number of obs   =       624
Group variable: panel                          Number of groups =        6

R-sq:  within = 0.4124                        obs per group: min =       104
        between = 0.9984                      avg =       104.0
        overall = 0.9731                      max =       104
```

```
corr(u_i, xb) = 0.9754                        F(5,5)          =      .
                                                Prob > F        =      .
```

(Std. Err. adjusted for 6 clusters in panel)

logproporc~n	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
retardo1lo~n	.4981634	.073581	6.77	0.001	.3090175	.6873093
retardo4lo~n	.1569713	.0670535	2.34	0.066	-.0153953	.3293378
x9	.006346	.0025935	2.45	0.058	-.0003208	.0130127
x10	.0072708	.0008382	8.67	0.000	.0051161	.0094254
x11	.0044697	.0008105	5.52	0.003	.0023863	.0065531
x24	-.0027524	.0011512	-2.39	0.062	-.0057115	.0002068
x26	-.0010845	.000176	-6.16	0.002	-.0015369	-.0006321
x27	-.0011324	.0004035	-2.81	0.038	-.0021696	-.0000951
x32	-.001221	.0002502	-4.88	0.005	-.0018641	-.0005779
x33	-.0015896	.0002555	-6.22	0.002	-.0022464	-.0009328
_cons	-1.023057	.040973	-24.97	0.000	-1.128382	-.9177328
sigma_u	.47651382					
sigma_e	.21073385					
rho	.83641628	(fraction of variance due to u_i)				



## 7.- CONCLUSIONES

Cuando comenzaba esta tesis tenía en mente dos grandes objetivos.

El primero, demostrar que era posible explicar lo que sucede en el punto de venta, y no sólo de una forma descriptiva, sino de una forma interactiva, esto es, determinando qué variables influyen y cuáles no, y el peso de esta influencia.

El segundo objetivo era demostrar empíricamente algo que las empresas llevan haciendo bastante tiempo en su práctica comercial. Esto es, demostrar que no existe una única política comercial para un producto en la empresa, sino que la política comercial es distinta según los formatos y los clientes.

El primer objetivo supuso crear un modelo de marketing, con la ayuda de técnicas estadísticas y econométricas, tomando como variable a explicar la cuota en tienda y como variables explicativas, todas aquellas variables, tanto de la empresa como de la competencia, que puedan influir sobre ella.

Dentro de los posibles modelos a utilizar, sin embargo se optó por construir un modelo con metodología de datos de panel. ¿Por qué este tipo de modelos? La razón de este tipo de modelos viene de considerar las circunstancias siguientes:

- Por un lado, con el tamaño de muestra con el que se trabaja, se hubiera podido estimar un modelo para cada formato y establecimiento. Este planteamiento nos hubiera llevado a estimar 24 modelos diferentes. Desde el punto de vista académico parece abordable, ahora bien, desde el punto de vista profesional, la dimensión del trabajo resulta un tanto exagerada, sobre todo si tenemos en cuenta que cada dos semanas sería necesario retroalimentar los modelos y actualizarlos. Ahora bien, y en un futuro como ampliación de investigación si cabe la posibilidad de trabajar en conjunto con todos los formatos utilizando la metodología de modelos SUR (regresiones aparentemente no relacionadas).

- En el otro extremo, se estimó un modelo único para todos los formatos y establecimientos. Este planteamiento nos hubiera llevado a demostrar que existe una única política comercial independientemente del formato comercializado o el punto de venta en el que se pone a disposición del consumidor. Sin embargo los contrastes de agrupabilidad que se realizaron echaron por tierra esta hipótesis, de tal manera que no podemos afirmar que exista una única política comercial.
- La última posibilidad que quedaba era buscar una metodología que permitiera agrupar, siempre que hubiera algún grado de homogeneidad, a los formatos y a los establecimientos. Esta metodología resultó ser los datos de panel, que imponiendo una restricción en la estimación (los parámetros son constantes entre los individuos del panel y las diferencias quedan recogidas en la constante y en la perturbación aleatoria), permitió la estimación de modelos por formatos y un modelo por tipo de establecimiento. Esta circunstancia está mucho más cerca del día a día de las empresas.

Respecto al segundo objetivo, una política comercial única por producto implica que no existe un tratamiento diferenciado para los distintos formatos que constituyen el producto. Nada más lejos de la realidad, los modelos planteados nos indican que las variables explicativas por formato son distintas. Esto quiere decir que la política comercial de la empresa varía según sea el formato e incluso el establecimiento, ya que los modelos también ponen de manifiesto que los pesos de las variables son distintos según sea el tipo de establecimiento considerado.

Veamos primeramente que variable han resultado ser significativas en los distintos modelos, independientemente del peso o influencia de cada una de ellas. En un primer momento trabajamos con los modelos iniciales para en un segundo momento, complicarlo y trabajar con los modelos finales.

En los modelos lineales, sin pasar el contraste de linealidad, las variables significativas son:

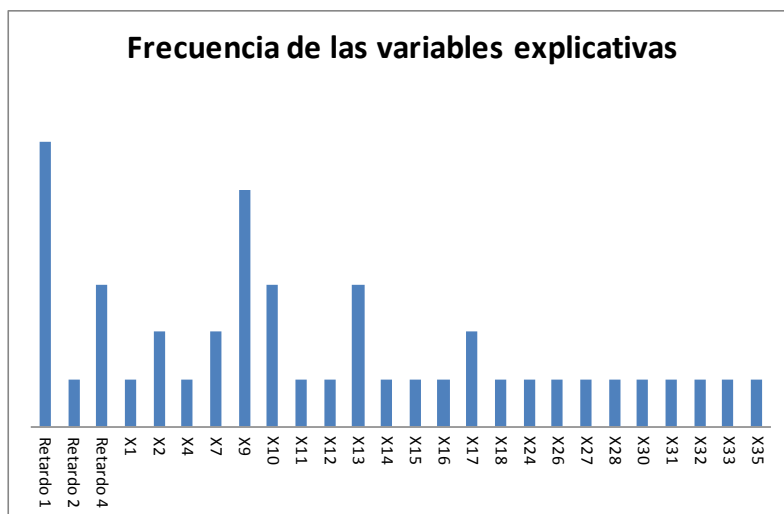
	Hasta 350 g.	De 350 a 499 g	De 500 a 899 g	De 900 a 1199 g	De 1200 a 1499 g	Hipermercados
Retardo 1	X	X	X	X	X	X
Retardo 2	X					
Retardo 4		X			X	X
X <sub>1</sub>				X		
X <sub>2</sub>		X		X		
X <sub>4</sub>			X			
X <sub>7</sub>	X	X				
X <sub>9</sub>		X	X	X	X	X
X <sub>10</sub>		X	X			X
X <sub>11</sub>						X
X <sub>12</sub>		X				
X <sub>13</sub>	X			X	X	
X <sub>14</sub>				X		
X <sub>15</sub>			X			
X <sub>16</sub>					X	
X <sub>17</sub>				X	X	
X <sub>18</sub>				X		
X <sub>24</sub>						X
X <sub>26</sub>						X
X <sub>27</sub>						X
X <sub>28</sub>				X		
X <sub>30</sub>		X				
X <sub>31</sub>	X					
X <sub>32</sub>						X
X <sub>33</sub>						X
X <sub>35</sub>				X		

Ordenadas según su frecuencia, las variables que han resultado ser significativas en los distintos modelos son:

- **Retardo 1.** La historia más reciente resulta ser influyente en todos los modelos considerados, por tanto el pasado está condicionando el futuro, si bien es cierto que este condicionamiento sólo es de un periodo.
- **X<sub>9</sub> Cobertura Promoción Folleto y Exposición Especial Marca 1.** Esta es la variable que aparece con mayor número de influencias. Sólo en el formato de hasta 350 gr. no aparece, esto último es debido a las especiales características de este formato, ya que el dominio de la Marca 1 es aplastante, no teniendo prácticamente ninguna competencia.

- ***Retardo 4,  $X_{10}$  Cobertura Promoción Sólo Exposición Especial Marca 1 y  $X_{13}$  Cobertura Reducción Temporal De Precio Marca 1.*** En este epígrafe aparecen tres variables con el mismo número de menciones; el retardo 4, de nuevo la historia no tan reciente, acaba influyendo en el presente. La variable  $X_{10}$  vuelve a poner de manifiesto la influencia de las acciones de merchandising realizadas en el punto de venta y por último la variable  $x_{13}$ , estas variables son variables de acción, variables controladas por la Marca 1, que hacen referencia a un marketing ofensivo, lo que hace la marca, frente a un marketing pasivo, lo que hace la competencia.
- ***$X_2$  Diferencial Precio Principal Competidor,  $X_7$  Distribución Ponderada Marcas de la Distribución,  $X_{17}$  Cobertura Promoción Sólo Exposición Especial Principal Competidor.*** Estas variables aparecen dos veces como significativas. De nuevo observamos una variable relacionada con el tema precio:  $X_2$ , y las restantes variables están relacionadas con otras variables distintas del marketing;  $X_7$  hace referencia a la distribución del producto, concretamente con las marcas de la distribución y  $X_{17}$  a la influencia de la acción de exposición especial de la competencia. Todas son variables manejadas por empresas competidoras, de manera que la Marca 1 sólo puede sufrir sus consecuencias. Se trata de variables de marketing pasivo.
- Por último aparecen un grupo de variables que sólo han sido significativas en cada uno de los modelos. Con la excepción de la variable ***Retardo 2*** para que este conjunto de información sea más manejable vamos a clasificarlas atendiendo a los criterios siguientes:

	Precio	Distribución	Promoción
Marca 1		$X_1$ Distribución Ponderada Marca 1	$X_{11}$ Cobertura Promoción Solo Folleto Marca 1
			$X_{12}$ Cobertura Otras Promociones Marca 1
			$X_{14}$ Cobertura Regalo Directo Marca 1
Principal Competidor	$X_{15}$ Intensidad De La Promoción De Precio Del Principal Competidor		$X_{18}$ Cobertura Promocion Solo Folleto Principal Competidor
	$X_{16}$ Cobertura Promocion Folleto Y Exposicion Especial Principal Competidor		
Competencia Resto de fabricantes	$X_{27}$ Cobertura Reducción Temporal De Precios Resto De Fabricantes		$X_{28}$ Cobertura Regalo Directo Resto De Fabricantes
			$X_{24}$ Cobertura Promoción Sólo Exposición Especial Resto De Fabricantes
			$X_{26}$ Cobertura Otras Promociones Resto De Fabricantes
Competencia Marcas de la Distribución	$X_4$ Diferencial Precio Marcas Distribución		$X_{30}$ Cobertura Promoción Folleto Y Exposición Especial Marcas Distribución
			$X_{31}$ Cobertura Promoción Solo Exposición Especial Marcas Distribución
			$X_{32}$ Cobertura Promoción Solo Folleto Marcas Distribución
			$X_{33}$ Cobertura Otras Promociones Marcas Distribución
			$X_{35}$ Cobertura Regalo Directo Marcas Distribución



La conclusión resulta obvia; las variables que influyen a la hora de explicar la cuota de mercado de una marca son distintas según se trata del formato considerado. Si entendemos que un producto está compuesto por un conjunto de formatos, podemos afirmar que no existe un tratamiento único por producto sino un tratamiento específico para cada formato. El conjunto de estos tratamientos forma lo que se llama la política comercial del producto.

Aún así, como se recordará, en capítulo anterior ha habido ocasiones donde el modelo lineal no ha pasado el contraste RESET de No linealidad, indicando que había que modificar la especificación del modelo, incluyendo algún componente no lineal que hasta ese momento no se había tenido en cuenta. Por lo tanto, se repite este análisis pero con las nuevas especificaciones de la ecuación de regresión.

	Hasta 350 g.	De 350-499 g	De 500-899 g	De 900-1199 g	De 1200-1499 g	Hipermercados
Retardo 1	X	X	X	X		X
Retardo 2	X					
Retardo 4		X			X	X
X <sub>1</sub>				X		
X <sub>2</sub>		X		X		
X <sub>4</sub>			X			
X <sub>7</sub>	X	X				
X <sub>9</sub>			X	X	X	X
X <sub>10</sub>		X	X			X
X <sub>11</sub>						X
X <sub>12</sub>		X				
X <sub>13</sub>	X			X	X	
X <sub>14</sub>				X		
X <sub>15</sub>			X			
X <sub>16</sub>					X	
X <sub>17</sub>				X	X	
X <sub>18</sub>				X		
X <sub>24</sub>						X
X <sub>26</sub>						X
X <sub>27</sub>						X
X <sub>28</sub>				X		
X <sub>30</sub>		X				
X <sub>31</sub>	X					
X <sub>32</sub>						X
X <sub>33</sub>						X
X <sub>35</sub>				X		
X <sub>9</sub> <sup>2</sup>		X				
X <sub>7</sub> X <sub>9</sub>		X				
X <sub>7</sub> X <sub>12</sub>		X				
Retardo 1 <sup>2</sup>					X	
Retardo 4 <sup>3</sup>		X			X	
X <sub>16</sub> <sup>2</sup>					X	
X <sub>2</sub> <sup>3</sup>		X				
X <sub>9</sub> <sup>3</sup>		X				
X <sub>13</sub> <sup>3</sup>					X	
X <sub>16</sub> <sup>3</sup>					X	
X <sub>18</sub> <sup>3</sup>				X		
X <sub>35</sub> <sup>3</sup>				X		

De este cuadro se extraen directamente tres conclusiones:

- Hay un grupo de modelos donde la forma funcional no presenta ningún tipo de No linealidad; estos son los modelos de Hasta 350 g e Hipermercados.
- Existe otro grupo de modelos donde la forma funcional presentaba problemas de No linealidad, y una vez corregida se ha obtenido una nueva especificación del modelo. Estos son los casos de los modelos de 350 a 499 g y de 1200 a 1499 g.
- Por último existe un modelo que presentando problemas de No linealidad, me ha resultado imposible encontrar la forma funcional correcta, y he preferido dejarlo con la ecuación original antes que complicarlos, en base al principio de parsimonia.

Una simple ojeada al cuadro pone de manifiesto que nuevamente las variables difieren entre los distintos modelos y por lo tanto estamos ante modelos distintos para los distintos formatos, reforzando aún más la idea de No existencia de una política comercial única para un producto.

## 7.1.- ¿QUÉ VARIABLES RESULTAN MÁS INFLUYENTES EN LOS DISTINTOS MODELOS?

La contestación a esta cuestión lleva aparejada unas consecuencias de acción importantes desde el ámbito de la política comercial de la marca. Esto es, parece más lógico centrar los esfuerzos de las empresas en aquellas variables que produzcan un mayor resultado en sus ventas.

El problema es cómo medir la influencia de dichas variables y comparar sus resultados, ya que cada una de ellas, ha sido medida en una unidad distinta y por lo tanto los resultados no son directamente comparables. La solución a este problema viene de la mano de un concepto absolutamente económico: La Elasticidad. El camino a seguir parece sencillo; calculamos la elasticidad para cada una de las variables explicativas y posteriormente las comparamos; aquellas variables con una elasticidad más elevada, influirán más sobre la variable Y.

El primer paso será definir claramente la elasticidad en un modelo lineal; es importante precisar el modelo de relación entre las variables porque como veremos más adelante el cálculo de la elasticidad está influido por la relación existente entre las variables.

$$y = \beta_1 + \beta_2 x + u$$

$$\varepsilon = \frac{\frac{dy}{y}}{\frac{dx}{x}} = \frac{dy}{dx} \cdot \frac{x}{y} = \beta_2 \cdot \frac{x}{y}$$

Este es el modelo, pero es desconocido, así que debemos estimarlo. De manera que la expresión cambia

$$\hat{y}_0 = b_1 + b_2 x_0$$

$$\hat{\varepsilon}_0 = b_2 \cdot \frac{x_0}{\hat{y}_0} = b_2 \cdot \frac{x_0}{b_1 + b_2 x_0}$$

Los cambios vienen producidos porque los parámetros han sido sustituidos por sus estimaciones y porque el cálculo de la elasticidad está particularizado para un punto  $x_0$  de la muestra. También es posible calcular la elasticidad en la media, en este caso los valores puntuales de las variables se sustituyen por sus valores medios

$$\hat{\varepsilon} = b_2 \cdot \frac{\bar{x}}{\bar{y}}$$

Una exposición detallada puede encontrarse en Lee C. Adkins y R. Carter Hill (2011, 62-84).



Otra cuestión a tratar es que la elasticidad obtenida, es una estimación de la elasticidad y por lo tanto, como cualquier estimación es necesario calcular su error standard, pero en este caso las reglas habituales de la varianza no pueden aplicarse ya que la relación existente no es lineal:  $b_2 \cdot \frac{x_0}{b_1 + b_2 x_0}$  y es preciso utilizar el método Delta.

### 7.1.1.- Elasticidad en el modelo Hasta 350 g.

El modelo especificado en este formato presenta la siguiente forma funcional:

$$\ln\left(\frac{p}{1-p}\right)_{it} = \delta + \theta_1 \ln\left(\frac{p}{1-p}\right)_{it-1} + \theta_2 \ln\left(\frac{p}{1-p}\right)_{it-2} + \beta_7 x_{7it} + \beta_{13} x_{13it} + \beta_{31} x_{31it} + u_{it}$$

Siendo:

- $\ln\left(\frac{p}{1-p}\right)_{it}$  la variable a explicar por el modelo
- $X_7$  Distribución Ponderada Marcas de la Distribución
- $X_{13}$  Cobertura de la Promoción sólo Exposición Especial de la Marca 1
- $X_{31}$  Cobertura Promoción Solo Exposición Especial Marcas Distribución

Para simplificar la notación haremos

$$\left(\frac{p}{1-p}\right)_{it} = y_{it}; \left(\frac{p}{1-p}\right)_{it-1} = y_{it-1}; \left(\frac{p}{1-p}\right)_{it-2} = y_{it-2}$$

De tal manera que el modelo queda escrito como:

$$\ln y_{it} = \delta + \theta_1 \ln y_{it-1} + \theta_2 \ln y_{it-2} + \beta_7 x_{7it} + \beta_{13} x_{13it} + \beta_{31} x_{31it} + u_{it}$$

Puede apreciarse, que este modelo es un modelo autorregresivo de retardos distribuidos, más concretamente, se trata de un ARDL (2,0) ya que la variable Y

aparece en el lado derecho de la regresión, con un retardo  $p=2$  y las variables  $X$  están todas ellas referidas al mismo momento del tiempo, por tanto  $q=0$ .

El modelo estimado para el formato de 350 gramos es.

$$\hat{Ln\hat{y}}_{it} = -2.410417 + 0.3137212\hat{Ln\hat{y}}_{it-1} + 0.1831895\hat{Ln\hat{y}}_{it-2} + 0.0040586x_{7it} + 0.0056823x_{13it} - 0.0051854x_{31it}$$

Es el momento de calcular las elasticidades del modelo y comparar los efectos de las variables.

En el proceso de cálculo, obtendremos primeramente los efectos marginales, para después calcular las elasticidades con las variables del modelo. Posteriormente se realizará la estimación de las elasticidades para cada variable y por tanto, la comparación de sus efectos.

La estimación de las elasticidades la realizaremos en la media de las variables y de esta manera podremos comparar sus efectos.

### Efectos marginales en el periodo t:

El efecto marginal se interpreta como la variación relativa ante un cambio unitario en la variable explicativa

$$\begin{aligned} \frac{\partial Ln\hat{y}_{it}}{\partial x_{7it}} &= \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{7it}}{x_{7it}}} = \beta_7 & ; \quad \frac{\partial y_{it}}{\partial x_{7it}} &= \beta_7 y_{it} \\ \frac{\partial Ln\hat{y}_{it}}{\partial x_{13it}} &= \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{13it}}{x_{13it}}} = \beta_{13} & ; \quad \frac{\partial y_{it}}{\partial x_{13it}} &= \beta_{13} y_{it} \\ \frac{\partial Ln\hat{y}_{it}}{\partial x_{31it}} &= \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{31it}}{x_{31it}}} = \beta_{31} & ; \quad \frac{\partial y_{it}}{\partial x_{31it}} &= \beta_{31} y_{it} \end{aligned}$$

$$\frac{\partial Lny_{it}}{\partial Lny_{it-1}} = \frac{\frac{\partial y_{it}}{y_{it-1}}}{\frac{\partial y_{it-1}}{y_{it-1}}} = \theta_1 ; \frac{\partial y_{it}}{\partial y_{it-1}} = \theta_1 \frac{y_{it}}{y_{it-1}}$$

$$\frac{\partial Lny_{it}}{\partial Lny_{it-2}} = \frac{\frac{\partial y_{it}}{y_{it-2}}}{\frac{\partial y_{it-2}}{y_{it-2}}} = \theta_2 ; \frac{\partial y_{it}}{\partial y_{it-2}} = \theta_1 \frac{y_{it}}{y_{it-2}}$$

### Elasticidad en el modelo:

$$\varepsilon_{x_{7it}} = \frac{\partial y_{it}}{\partial x_{7it}} \frac{x_{7it}}{y_{it}} ; \varepsilon_{x_{7it}} = \beta_7 y_{it} \frac{x_{7it}}{y_{it}} ; \varepsilon_{x_{7it}} = \beta_7 x_{7it}$$

$$\varepsilon_{x_{13it}} = \frac{\partial y_{it}}{\partial x_{13it}} \frac{x_{13it}}{y_{it}} ; \varepsilon_{x_{13it}} = \beta_{13} y_{it} \frac{x_{13it}}{y_{it}} ; \varepsilon_{x_{13it}} = \beta_{13} x_{13it}$$

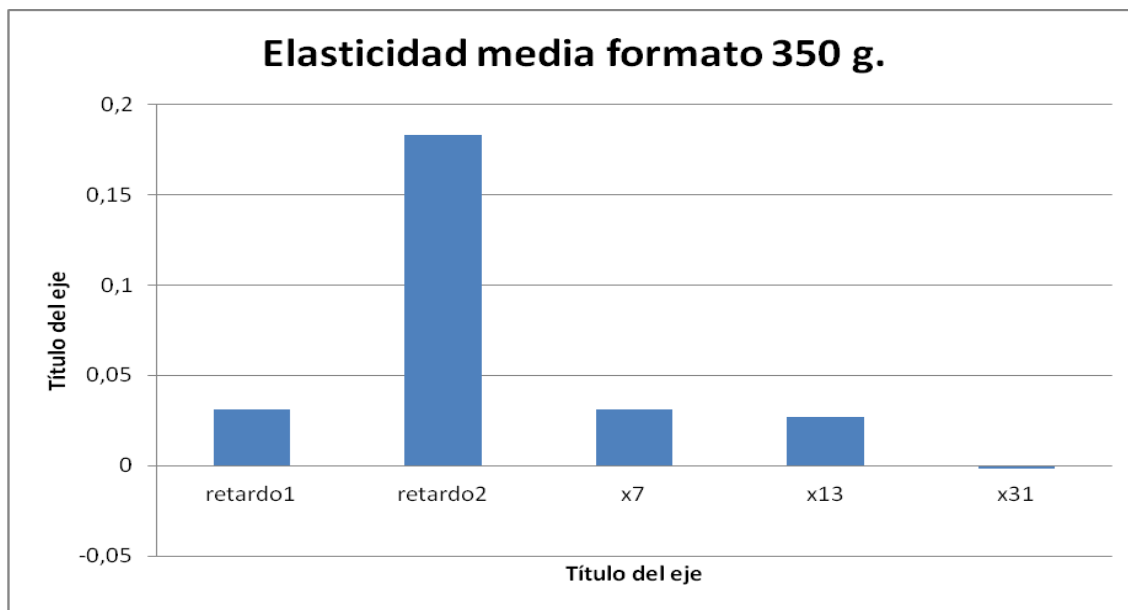
$$\varepsilon_{x_{31it}} = \frac{\partial y_{it}}{\partial x_{31it}} \frac{x_{31it}}{y_{it}} ; \varepsilon_{x_{31it}} = \beta_{31} y_{it} \frac{x_{31it}}{y_{it}} ; \varepsilon_{x_{31it}} = \beta_{31} x_{31it}$$

$$\varepsilon_{y_{it-1}} = \frac{\partial y_{it}}{\partial y_{it-1}} \frac{y_{it-1}}{y_{it}} ; \varepsilon_{y_{it-1}} = \theta_1 \frac{y_{it}}{y_{it-1}} \frac{y_{it-1}}{y_{it}} ; \varepsilon_{y_{it-1}} = \theta_1$$

$$\varepsilon_{y_{it-2}} = \frac{\partial y_{it}}{\partial y_{it-2}} \frac{y_{it-2}}{y_{it}} ; \varepsilon_{y_{it-2}} = \theta_2 \frac{y_{it}}{y_{it-2}} \frac{y_{it-2}}{y_{it}} ; \varepsilon_{y_{it-2}} = \theta_2$$

La estimación de la elasticidad para el promedio de las variables es:

Variables	Elasticidad en los establecimientos				
	Elasticidad en la media	Hiper	Super de 1000 a 2499 m	Super de 400 a 999 m	Super de 100 a 399 m
<b>Yit-1</b>	0,03137212	0,03137212	0,03137212	0,03137212	0,03137212
<b>Yit-2</b>	0,1831895	0,183190	0,183190	0,183190	0,183190
<b>x7</b>	0,03128185	0,050349612	0,03897788	0,02818047	0,00761945
<b>x13</b>	0,02682399	0,030694182	0,02484564	0,02685406	0,02490208
<b>x31</b>	-0,00145391	-0,00255438	0	0	-0,00326126



El cuadro y el gráfico ponen de manifiesto que la variable que más influye en la cuota de mercado de la marca es el ***retardo de hace dos periodos***.

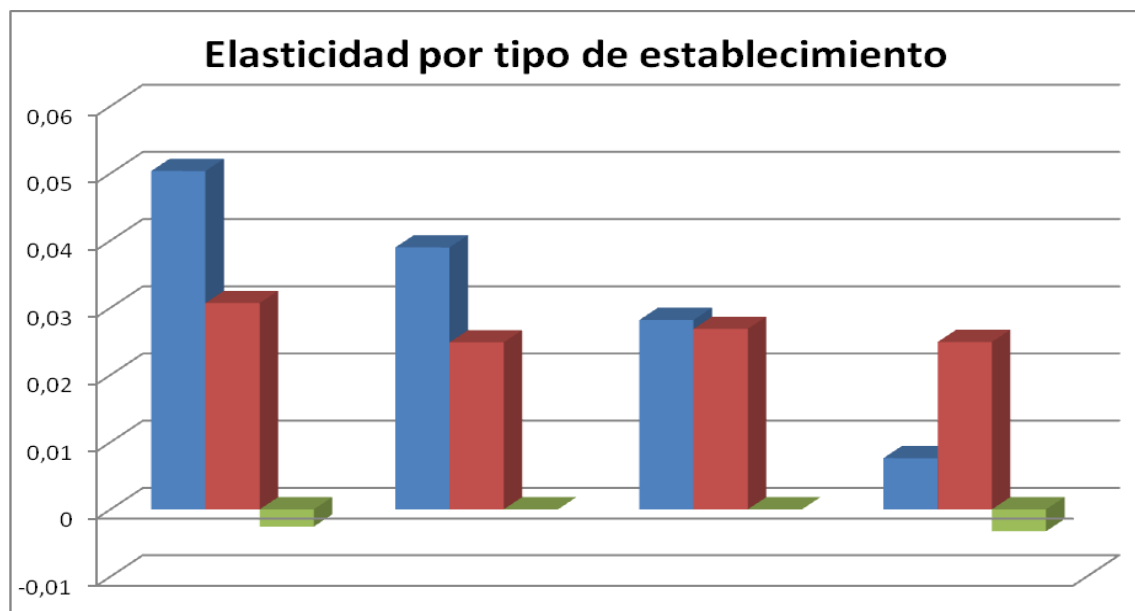
En un segundo lugar aparecen; con prácticamente la misma importancia, las variables ***retardo del último periodo***,  $x_7$ : ***Distribución Ponderada Marcas de la Distribución*** y  $x_{13}$ : ***Cobertura Reducción Temporal De Precio Marca 1***, es este tipo de acción comercial la que consigue en este formato movilizar la cuota de mercado de la marca.

Por último está la variable  $x_{31}$ : ***Cobertura Promoción Solo Exposición Especial Marcas Distribución***, de manera que son la intensidad de las acciones de las marcas de la distribución en promociones de exposición especial las que reducen la cuota en tienda de la Marca 1.

Ahora bien, si estudiamos las elasticidades por tipo de establecimiento, es cierto que el criterio general es el mismo, pero la importancia que tiene cada variable varía según el tipo de establecimiento considerado.

En el gráfico no se ha incluido las variables retardo 1 y retardo 2 porque la elasticidad es constante. Sólo se incluyen aquellas variables:  $x_7$ ,  $x_{13}$  y  $x_{31}$  que

muestran valores distintos. Esto es posible gracias a que la elasticidad no es constante y depende tanto del periodo como del establecimiento considerado.



### 7.1.2.- Elasticidad en el modelo de 350 A 499 g.

Cuando hablamos del formato de 350 a 499 g. el modelo definitivo viene definido por la siguiente forma funcional:

$$\begin{aligned} \ln\left(\frac{p}{1-p}\right)_{it} = & \delta + \theta_1 \ln\left(\frac{p}{1-p}\right)_{it-1} + \theta_4 \ln\left(\frac{p}{1-p}\right)_{it-4} + \beta_2 x_{2it} + \beta_7 x_{7it} + \beta_{10} x_{10it} + \beta_{30} x_{30it} + \beta_{209} x_{9it}^2 + \beta_{79} x_{7it} x_{9it} \\ & + \beta_{712} x_{7it} x_{12it} + \theta_{300} \ln\left(\frac{p}{1-p}\right)_{it-4}^3 + \beta_{302} x_{2it}^3 + \beta_{309} x_{9it}^3 + u_{it} \end{aligned}$$

Nuevamente hacemos un cambio en la notación  $\left(\frac{p}{1-p}\right) = y$  para que su lectura sea más sencilla. El modelo queda de la siguiente manera.

$$\begin{aligned} \ln y_{it} = & \delta + \theta_1 \ln y_{it-1} + \theta_4 \ln y_{it-4} + \beta_2 x_{2it} + \beta_7 x_{7it} + \beta_{10} x_{10it} + \beta_{30} x_{30it} + \beta_{209} x_{9it}^2 + \beta_{79} x_{7it} x_{9it} \\ & + \beta_{712} x_{7it} x_{12it} + \theta_{300} \ln y_{it-4}^3 + \beta_{302} x_{2it}^3 + \beta_{309} x_{9it}^3 + u_{it} \end{aligned}$$

El camino a seguir es idéntico al formato anterior. Calculamos los efectos marginales y a continuación las elasticidades de cada variable.

#### Efectos marginales en el periodo t:

El efecto marginal se interpreta como la variación relativa ante un cambio unitario en la variable explicativa

$$\frac{\partial \ln y_{it}}{\partial x_{2it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{2it}}{x_{2it}}} = \beta_2 + 3\beta_{302} x_{2it}^2 ; \frac{\partial y_{it}}{\partial x_{2it}} = (\beta_2 + 3\beta_{302} x_{2it}^2) y_{it}$$

$$\frac{\partial \ln y_{it}}{\partial x_{7it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{7it}}{x_{7it}}} = \beta_7 + \beta_{79} x_{9it} + \beta_{712} x_{12it} ; \frac{\partial y_{it}}{\partial x_{7it}} = (\beta_7 + \beta_{79} x_{9it} + \beta_{712} x_{12it}) y_{it}$$

$$\frac{\partial \ln y_{it}}{\partial x_{9it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{9it}}{x_{9it}}} = 2\beta_{209} x_{9it} + \beta_{79} x_{7it} + 3\beta_{309} x_{9it}^2 ; \frac{\partial y_{it}}{\partial x_{9it}} = (2\beta_{209} x_{9it} + \beta_{79} x_{7it} + 3\beta_{309} x_{9it}^2) y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{10it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial y_{it}}{\partial x_{10it}}} = \beta_{10} ; \frac{\partial y_{it}}{\partial x_{10it}} = \beta_{10} y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{12it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial y_{it}}{\partial x_{12it}}} = \beta_{712} x_{7it} ; \frac{\partial y_{it}}{\partial x_{12it}} = (\beta_{712} x_{7it}) y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{30it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial y_{it}}{\partial x_{30it}}} = \beta_{30} ; \frac{\partial y_{it}}{\partial x_{30it}} = \beta_{30} y_{it}$$

$$\frac{\partial Lny_{it}}{\partial Lny_{it-1}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial y_{it-1}}{y_{it-1}}} = \theta_1 ; \frac{\partial y_{it}}{\partial y_{it-1}} = \theta_1 \frac{y_{it}}{y_{it-1}}$$

$$\frac{\partial Lny_{it}}{\partial Lny_{it-4}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial y_{it-4}}{y_{it-4}}} = \theta_4 + 3\theta_{304} \frac{y_{it-4}^2}{y_{it-4}^3} ; \frac{\partial y_{it}}{\partial y_{it-4}} = \left( \theta_4 + 3\theta_{304} \frac{y_{it-4}^2}{y_{it-4}^3} \right) \frac{y_{it}}{y_{it-4}}$$

## Elasticidad

$$\varepsilon_{x_{2it}} = \frac{\partial y_{it}}{\partial x_{2it}} \frac{x_{2it}}{y_{it}} ; \varepsilon_{x_{2it}} = (\beta_2 + 3\beta_{302} x_{2it}^2) y_{it} \frac{x_{2it}}{y_{it}} ; \varepsilon_{x_{2it}} = (\beta_2 + 3\beta_{302} x_{2it}^2) x_{2it}$$

$$\varepsilon_{x_{2it}} = \beta_2 x_{2it} + 3\beta_{302} x_{2it}^3$$

$$\varepsilon_{x_{7it}} = \frac{\partial y_{it}}{\partial x_{7it}} \frac{x_{7it}}{y_{it}} ; \varepsilon_{x_{7it}} = (\beta_7 + \beta_{79} x_{9it} + \beta_{712} x_{12it}) y_{it} \frac{x_{7it}}{y_{it}} ; \varepsilon_{x_{7it}} = (\beta_7 + \beta_{79} x_{9it} + \beta_{712} x_{12it}) y_{it} \frac{x_{7it}}{y_{it}} ;$$

$$\varepsilon_{x_{7it}} = (\beta_7 + \beta_{79} x_{9it} + \beta_{712} x_{12it}) x_{7it} ; \varepsilon_{x_{7it}} = \beta_7 x_{7it} + \beta_{79} x_{7it} x_{9it} + \beta_{712} x_{7it} x_{12it}$$

$$\varepsilon_{x_{9it}} = \frac{\partial y_{it}}{\partial x_{9it}} \frac{x_{9it}}{y_{it}} ; \varepsilon_{x_{9it}} = (2\beta_{209} x_{9it} + \beta_{79} x_{7it} + 3\beta_{309} x_{9it}^2) y_{it} \frac{x_{9it}}{y_{it}} ;$$

$$\varepsilon_{x_{9it}} = (2\beta_{209} x_{9it} + \beta_{79} x_{7it} + 3\beta_{309} x_{9it}^2) x_{9it} ; \varepsilon_{x_{9it}} = 2\beta_{209} x_{9it}^2 + \beta_{79} x_{7it} x_{9it} + 3\beta_{309} x_{9it}^3$$

$$\varepsilon_{x_{10it}} = \frac{\partial y_{it}}{\partial x_{10it}} \frac{x_{10it}}{y_{it}} ; \varepsilon_{x_{10it}} = \beta_{10} y_{it} \frac{x_{10it}}{y_{it}} ; \varepsilon_{x_{10it}} = \beta_{10} x_{10it}$$

$$\varepsilon_{x_{12it}} = \frac{\partial y_{it}}{\partial x_{12it}} \frac{x_{12it}}{y_{it}} ; \varepsilon_{x_{12it}} = (\beta_{712} x_{7it}) y_{it} \frac{x_{12it}}{y_{it}} ; \varepsilon_{x_{12it}} = (\beta_{712} x_{7it}) y_{it} \frac{x_{12it}}{y_{it}} ; \varepsilon_{x_{12it}} = \beta_{712} x_{7it} x_{12it}$$

$$\varepsilon_{x_{30it}} = \frac{\partial y_{it}}{\partial x_{30it}} \frac{x_{30it}}{y_{it}} ; \varepsilon_{x_{30it}} = \beta_{30} y_{it} \frac{x_{30it}}{y_{it}} ; \varepsilon_{x_{30it}} = \beta_{30} x_{30it}$$

$$\varepsilon_{y_{it-1}} = \frac{\partial y_{it}}{\partial y_{it-1}} \frac{y_{it-1}}{y_{it}} ; \varepsilon_{y_{it-1}} = \theta_1 \frac{y_{it}}{y_{it-1}} \frac{y_{it-1}}{y_{it}} ; \varepsilon_{y_{it-1}} = \theta_1$$

$$\varepsilon_{y_{it-4}} = \frac{\partial y_{it}}{\partial y_{it-4}} \frac{y_{it-4}}{y_{it}} ; \varepsilon_{y_{it-4}} = \left( \theta_4 + 3\theta_{304} \frac{y_{it-4}^2}{y_{it-4}^3} \right) \frac{y_{it}}{y_{it-4}} \frac{y_{it-4}}{y_{it}} ; \varepsilon_{y_{it-4}} = \theta_4 + \frac{3\theta_{304}}{y_{it-4}}$$

Es el momento de estimar estas elasticidades, por tanto es preciso pasar al modelo estimado.

$$\begin{aligned} \ln \hat{y}_{it} = & 2.492593 + 0.3580636 \ln y_{it-1} - 0.817763 \ln y_{it-4} - 8.563382 x_{2it} - 0.0056555 x_{7it} \\ & + 0.0082547 x_{10it} - 0.0050432 x_{30it} - 0.0724968 x_{9it}^2 + 0.0065381 x_{7it} x_{9it} + 0.0005653 x_{7it} x_{12it} \\ & - 0.1627271 \ln y_{it-4}^3 + 3.506264 x_{2it}^3 + 0.0054587 x_{9it}^3 \end{aligned}$$

El cálculo de las elasticidades en la media y en los diferentes tipos de establecimientos se recoge en el cuadro que viene a continuación.

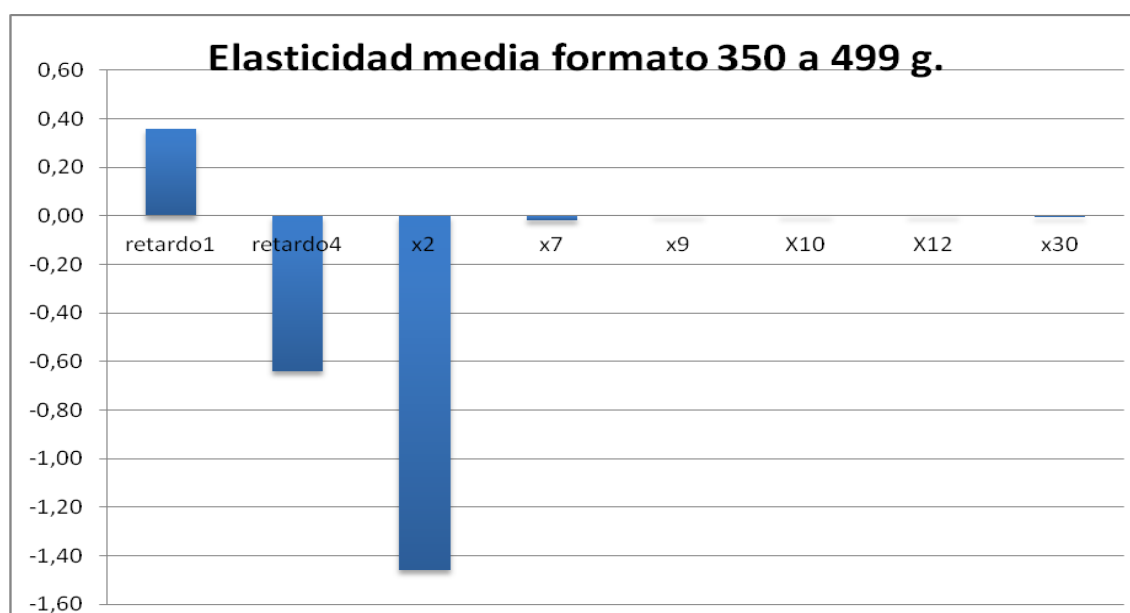
Variables	Elasticidad en los establecimientos				
	Elasticidad en la media	Hiper	Super de 1000 a 2499 m	Super de 400 a 999 m	Super de 100 a 399 m
<b>Yit-1</b>	0,3580636	0,3580636	0,3580636	0,3580636	0,3580636
<b>Yit-4</b>	-0,640023488	-0,657821075	-0,632294282	-0,647050301	-0,617884316
<b>x2</b>	-1,457748973	-0,910973006	-1,634526212	-1,612568774	-1,610320253
<b>x7</b>	-0,019484704	-0,351456693	-0,099797705	-0,124910427	-0,22386872
<b>x9</b>	0,000991864	0	0,034549013	0	0
<b>x10</b>	0,002243205	0	0,000992151	0	0,007980667
<b>x12</b>	4,23703E-05	0	0	0	0,001807504
<b>x30</b>	-0,001072866	-0,000381251	-0,003313243	0	-0,00059697

En el gráfico puede observarse la comparación de los efectos de las variables.

De tal manera que las variables que están relacionadas directamente con la cuota de ventas de la marca son la **cuota de ventas lograda en el periodo anterior** y las variables  $x_9$  **Cobertura Promoción Folleto y Exposición Especial Marca 1**,  $x_{10}$  **Cobertura Promoción Solo Exposición Especial Marca 1** y  $x_{12}$  **Cobertura Otras Promociones Marca 1**.



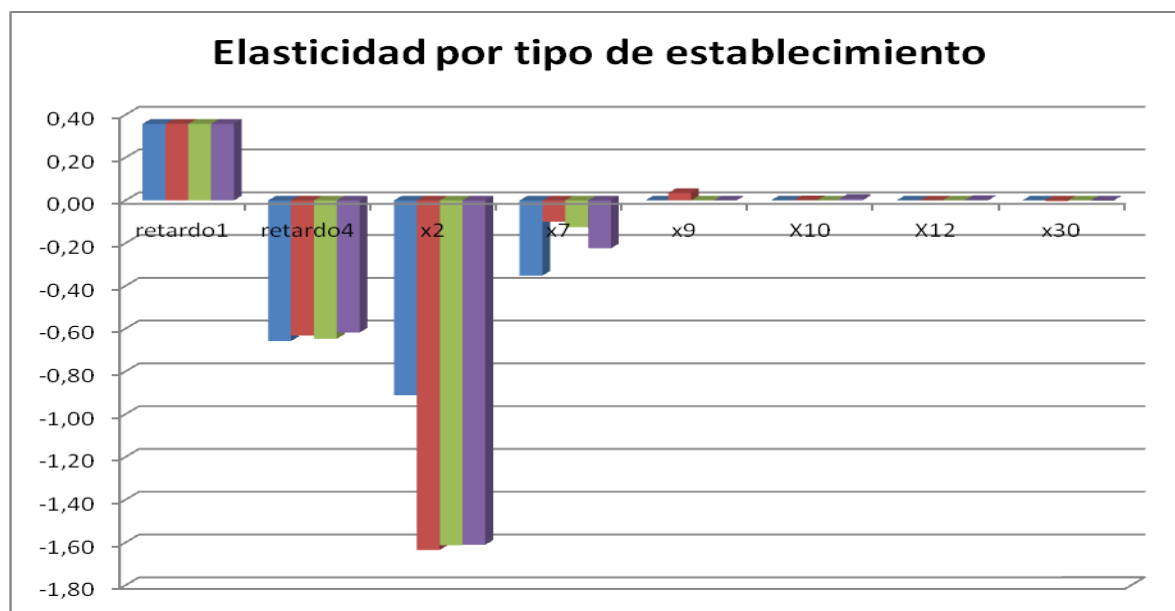
Las variables que están relacionadas inversamente con la cuota de mercado son la cuota de ventas lograda hace cuatro periodos  $x_2$  **Diferencial Precio Principal Competidor**,  $x_7$  **Distribución Ponderada Marcas Distribución**,  $x_{30}$  **Cobertura Promoción Folleto y Exposición Especial Marcas Distribución**. Por lo tanto puede afirmarse que las amenazas de la Marca 1, en este formato, viene del lado de las marcas de distribución y del principal competidor, pero no del Resto de Fabricantes.



Cuando se estudia la elasticidad media por tipo de establecimiento, los signos de las variables, son los mismos, pero cambia la intensidad con la que afecta al elasticidad a las variables.

Así, la variable **Diferencial Precio Principal Competidor**  $X_2$ , tiene un comportamiento curioso, afectando en mayor medida en los supermercados que en los hipermercados, es posible que sea como consecuencia de que el Principal Competidor está más presenta en los supermercados, mientras que en los hipermercados, su papel está ocupado por las marcas propias/ blancas de los establecimientos.

Otro tanto parece ocurrir con la variable ***Distribución Ponderada Marcas de la Distribución***  $X_7$  y que refuerza la conclusión anterior, y es que esta variable tiene especial influencia en el hipermercado, donde las marcas blancas son más abundantes, reduciendo su importancia en los supermercados y llamando la atención el papel que juega en los supermercados más pequeños.



### 7.1.3.- Elasticidad en el modelo de 500 a 899 g.

$$Lny_{it} = \delta + \theta_1 Lny_{it-1} + \beta_4 x_{4it} + \beta_9 x_{9it} + \beta_{10} x_{10it} + \beta_{15} x_{15it} + u_{it}$$

#### Efectos marginales

$$\begin{aligned} \frac{\partial Lny_{it}}{\partial x_{4it}} &= \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{4it}}{x_{4it}}} = \beta_4 \quad ; \quad \frac{\partial y_{it}}{\partial x_{4it}} = \beta_4 y_{it} \\ \frac{\partial Lny_{it}}{\partial x_{9it}} &= \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{9it}}{x_{9it}}} = \beta_9 \quad ; \quad \frac{\partial y_{it}}{\partial x_{9it}} = \beta_9 y_{it} \\ \frac{\partial Lny_{it}}{\partial x_{10it}} &= \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{10it}}{x_{10it}}} = \beta_{10} \quad ; \quad \frac{\partial y_{it}}{\partial x_{10it}} = \beta_{10} y_{it} \\ \frac{\partial Lny_{it}}{\partial x_{15it}} &= \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{15it}}{x_{15it}}} = \beta_{15} \quad ; \quad \frac{\partial y_{it}}{\partial x_{15it}} = \beta_{15} y_{it} \\ \frac{\partial Lny_{it}}{\partial Lny_{it-1}} &= \frac{\frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial y_{it-1}}{y_{it-1}}}}{\frac{\partial y_{it-1}}{y_{it-1}}} = \theta_1 \quad ; \quad \frac{\partial y_{it}}{\partial y_{it-1}} = \theta_1 \frac{y_{it}}{y_{it-1}} \end{aligned}$$

#### Elasticidad

$$\begin{aligned} \varepsilon_{x_{4it}} &= \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{4it}}{x_{4it}}} = \beta_4 \frac{x_{4it}}{y_{it}} \quad ; \quad \varepsilon_{x_{9it}} = \beta_9 \frac{x_{9it}}{y_{it}} \quad ; \quad \varepsilon_{x_{10it}} = \beta_{10} \frac{x_{10it}}{y_{it}} \\ \varepsilon_{x_{9it}} &= \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{9it}}{x_{9it}}} = \beta_9 \frac{x_{9it}}{y_{it}} \quad ; \quad \varepsilon_{x_{10it}} = \beta_{10} \frac{x_{10it}}{y_{it}} \quad ; \quad \varepsilon_{x_{15it}} = \beta_{15} \frac{x_{15it}}{y_{it}} \\ \varepsilon_{x_{10it}} &= \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{10it}}{x_{10it}}} = \beta_{10} \frac{x_{10it}}{y_{it}} \quad ; \quad \varepsilon_{x_{15it}} = \beta_{15} \frac{x_{15it}}{y_{it}} \quad ; \quad \varepsilon_{x_{15it}} = \beta_{15} \frac{x_{15it}}{y_{it}} \\ \varepsilon_{x_{15it}} &= \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{15it}}{x_{15it}}} = \beta_{15} \frac{x_{15it}}{y_{it}} \quad ; \quad \varepsilon_{x_{15it}} = \beta_{15} \frac{x_{15it}}{y_{it}} \quad ; \quad \varepsilon_{x_{15it}} = \beta_{15} \frac{x_{15it}}{y_{it}} \\ \varepsilon_{y_{it-1}} &= \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial y_{it-1}}{y_{it-1}}} = \theta_1 \frac{y_{it}}{y_{it-1}} \quad ; \quad \varepsilon_{y_{it-1}} = \theta_1 \frac{y_{it}}{y_{it-1}} \quad ; \quad \varepsilon_{y_{it-1}} = \theta_1 \end{aligned}$$

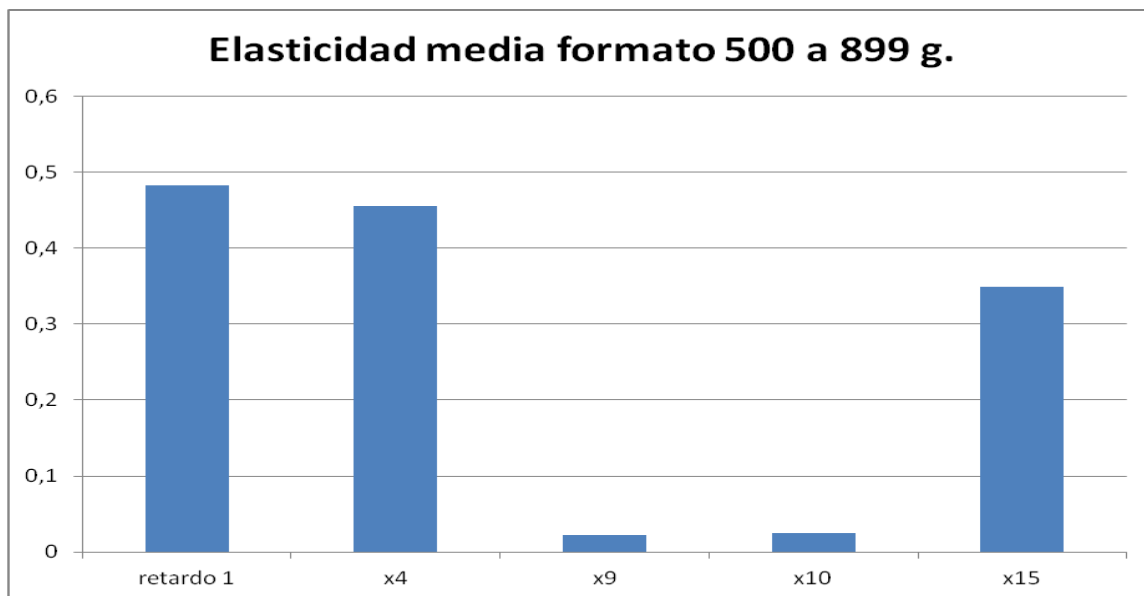
El modelo estimado viene dado por la estructura que viene a continuación y que nos permitirá estimar la elasticidad en este formato.

$$Lny_{it} = -1,742882 + 0,483284 Lny_{it-1} + 0,803612 x_{4it} + 0,0089608 x_{9it} + 0,0075775 x_{10it} + 0,3612625 x_{15it}$$

Variables	Elasticidad en la media	Elasticidad en los establecimientos			
		Hiper	Super de 1000 a 2499 m	Super de 400 a 999 m	Super de 100 a 399 m
Yit-1	0,483284	0,483284	0,483284	0,483284	0,483284
x4	0,45491583	0,47056658	0,44291244	0,45168427	0,45449996
x9	0,02248004	0,00523411	0,02342526	0,03053447	0,03072632
x10	0,02526111	0,00555849	0,02770454	0,03056391	0,03721751
x15	0,34921873	0,34891216	0,34771183	0,34707001	0,35318088

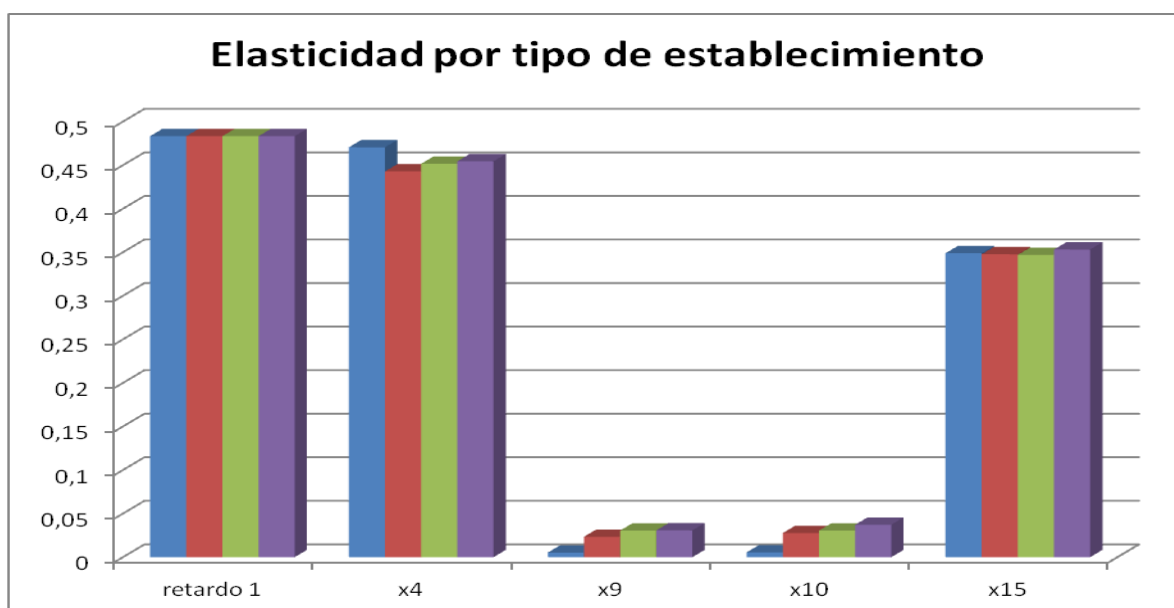
Las variables con un efecto más importante sobre la cuota de mercado son, por orden de importancia:

- **Primer retardo de la variable** estudiada, con signo positivo, por lo que actuaciones positivas en el pasado, influyen positivamente en el futuro, pero actuaciones negativas prolongan su influencia en el tiempo.
- **Diferencial Precio Marcas Distribución (X<sub>4</sub>)** Esta variable nos habla de la influencia del precio respecto a las Marcas de Distribución como elemento dinamizador de la cuota en tienda de la Marca 1.
- **Intensidad de la Promoción de Precio del Principal Competidor (X<sub>15</sub>)**. Aunque con signo diferente al esperado, aparece con una importancia elevada en su efecto sobre la cuota en tienda de la Marca 1.
- Las dos últimas posiciones están ocupadas por **Cobertura Promoción Sólo Exposición Especial Marca 1 (X<sub>10</sub>)** y **Cobertura Promoción Folleto y Exposición Especial Marca 1 (X<sub>9</sub>)** indicando en este caso que valores bajos de la variable (Cobertura de la promoción) se asocian a valores bajos de la cuota en tienda y valores elevados se asocian a valores altos de la cuota en tienda. Por tanto el signo de las variables parece lógico y adecuado.



Cuando analizamos los efectos de las variables sobre la cuota en tienda atendiendo a los distintos tipos de establecimientos se extraen las siguientes conclusiones.

- La variable retardada tiene exactamente la misma importancia en los distintos tipos de establecimientos, ya que se trata de una constante.
- En los supermercados más pequeños, los efectos de las variables son siempre mayores que en los restantes tipos de supermercados e incluso hiper; la única excepción es en la variable Diferencial Precio Marcas Distribución ( $X_4$ ) en la que no se cumple esta afirmación.



#### 7.1.4.- Elasticidad en el modelo de 900 A 1199 g.

$$Lny_{it} = \delta + \theta_1 Lny_{it-1} + \beta_1 x_{1it} + \beta_2 x_{2it} + \beta_9 x_{9it} + \beta_{13} x_{13it} + \beta_{14} x_{14it} + \beta_{17} x_{17it} + \beta_{18} x_{18it} \\ + \beta_{28} x_{28it} + \beta_{35} x_{35it} + \beta_{318} x_{18it}^3 + \beta_{335} x_{35it}^3 + u_{it}$$

#### Efectos marginales

$$\frac{\partial Lny_{it}}{\partial x_{1it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{1it}}{x_{1it}}} = \beta_1 \quad ; \quad \frac{\partial y_{it}}{\partial x_{1it}} = \beta_1 y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{2it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{2it}}{x_{2it}}} = \beta_2 \quad ; \quad \frac{\partial y_{it}}{\partial x_{2it}} = \beta_2 y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{9it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{9it}}{x_{9it}}} = \beta_9 \quad ; \quad \frac{\partial y_{it}}{\partial x_{9it}} = \beta_9 y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{13it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{13it}}{x_{13it}}} = \beta_{13} \quad ; \quad \frac{\partial y_{it}}{\partial x_{13it}} = \beta_{13} y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{14it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{14it}}{x_{14it}}} = \beta_{14} \quad ; \quad \frac{\partial y_{it}}{\partial x_{14it}} = \beta_{14} y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{13it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{13it}}{x_{13it}}} = \beta_{13} \quad ; \quad \frac{\partial y_{it}}{\partial x_{13it}} = \beta_{13} y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{17it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{17it}}{x_{17it}}} = \beta_{17} \quad ; \quad \frac{\partial y_{it}}{\partial x_{17it}} = \beta_{17} y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{18it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{18it}}{x_{18it}}} = \beta_{18} + 3\beta_{318} x_{18it}^2 \quad ; \quad \frac{\partial y_{it}}{\partial x_{18it}} = (\beta_{18} + 3\beta_{318} x_{18it}^2) y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{28it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{28it}}{x_{28it}}} = \beta_{28} \quad ; \quad \frac{\partial y_{it}}{\partial x_{28it}} = \beta_{28} y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{35it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{35it}}{x_{35it}}} = \beta_{35} + 3\beta_{335} x_{35it}^2 \quad ; \quad \frac{\partial y_{it}}{\partial x_{35it}} = (\beta_{35} + 3\beta_{335} x_{35it}^2) y_{it}$$

$$\frac{\partial Lny_{it}}{\partial Lny_{it-1}} = \frac{\frac{\partial y_{it}}{y_{it-1}}}{\frac{\partial y_{it-1}}{y_{it-1}}} = \theta_1 \quad ; \quad \frac{\partial y_{it}}{\partial y_{it-1}} = \theta_1 \frac{y_{it}}{y_{it-1}}$$

## Elasticidades

$$\begin{aligned} \varepsilon_{x_{1it}} &= \frac{\partial y_{it}}{\partial x_{1it}} \frac{x_{1it}}{y_{it}} \quad ; \varepsilon_{x_{1it}} = \beta_1 y_{it} \frac{x_{1it}}{y_{it}} \quad ; \varepsilon_{x_{1it}} = \beta_1 x_{1it} \\ \varepsilon_{x_{2it}} &= \frac{\partial y_{it}}{\partial x_{2it}} \frac{x_{2it}}{y_{it}} \quad ; \varepsilon_{x_{2it}} = \beta_2 y_{it} \frac{x_{2it}}{y_{it}} \quad ; \varepsilon_{x_{2it}} = \beta_2 x_{2it} \\ \varepsilon_{x_{9it}} &= \frac{\partial y_{it}}{\partial x_{9it}} \frac{x_{9it}}{y_{it}} \quad ; \varepsilon_{x_{9it}} = \beta_9 y_{it} \frac{x_{9it}}{y_{it}} \quad ; \varepsilon_{x_{9it}} = \beta_9 x_{9it} \\ \varepsilon_{x_{13it}} &= \frac{\partial y_{it}}{\partial x_{13it}} \frac{x_{13it}}{y_{it}} \quad ; \varepsilon_{x_{13it}} = \beta_{13} y_{it} \frac{x_{13it}}{y_{it}} \quad ; \varepsilon_{x_{13it}} = \beta_{13} x_{13it} \\ \varepsilon_{x_{14it}} &= \frac{\partial y_{it}}{\partial x_{14it}} \frac{x_{14it}}{y_{it}} \quad ; \varepsilon_{x_{14it}} = \beta_{14} y_{it} \frac{x_{14it}}{y_{it}} \quad ; \varepsilon_{x_{14it}} = \beta_{14} x_{14it} \\ \varepsilon_{x_{17it}} &= \frac{\partial y_{it}}{\partial x_{17it}} \frac{x_{17it}}{y_{it}} \quad ; \varepsilon_{x_{17it}} = \beta_{17} y_{it} \frac{x_{17it}}{y_{it}} \quad ; \varepsilon_{x_{17it}} = \beta_{17} x_{17it} \\ \varepsilon_{x_{18it}} &= \frac{\partial y_{it}}{\partial x_{18it}} \frac{x_{18it}}{y_{it}} \quad ; \varepsilon_{x_{18it}} = (\beta_{18} + 3\beta_{318} x_{18it}^2) y_{it} \frac{x_{18it}}{y_{it}} \quad ; \varepsilon_{x_{18it}} = (\beta_{18} + 3\beta_{318} x_{18it}^2) x_{18it} \quad ; \varepsilon_{x_{18it}} = \beta_{18} x_{18it} + 3\beta_{318} x_{18it}^3 \\ \varepsilon_{x_{28it}} &= \frac{\partial y_{it}}{\partial x_{28it}} \frac{x_{28it}}{y_{it}} \quad ; \varepsilon_{x_{28it}} = \beta_{28} y_{it} \frac{x_{28it}}{y_{it}} \quad ; \varepsilon_{x_{28it}} = \beta_{28} x_{28it} \\ \varepsilon_{x_{35it}} &= \frac{\partial y_{it}}{\partial x_{35it}} \frac{x_{35it}}{y_{it}} \quad ; \varepsilon_{x_{35it}} = (\beta_{35} + 3\beta_{335} x_{35it}^2) y_{it} \frac{x_{35it}}{y_{it}} \quad ; \varepsilon_{x_{35it}} = (\beta_{35} + 3\beta_{335} x_{35it}^2) x_{35it} \quad ; \varepsilon_{x_{35it}} = \beta_{35} x_{35it} + 3\beta_{335} x_{35it}^3 \\ \varepsilon_{y_{it-1}} &= \frac{\partial y_{it}}{\partial y_{it-1}} \frac{y_{it-1}}{y_{it}} \quad ; \varepsilon_{y_{it-1}} = \theta_1 \frac{y_{it}}{y_{it-1}} \frac{y_{it-1}}{y_{it}} \quad ; \varepsilon_{y_{it-1}} = \theta_1 \end{aligned}$$

La estructura estimada del modelo se presenta a continuación.

$$\begin{aligned} Ln\hat{y}_{it} &= -1,552604 + 0,3966162Ln y_{it-1} - 0,0027039x_{1it} + 0,5402461x_{2it} + 0,0098596x_{9it} + 0,0065626x_{13it} \\ &+ 0,0089146x_{14it} + 0,0013155x_{17it} + 0,0046501x_{18it} + 0,0547714x_{28it} + 0,4758598x_{35it} - 0,0000025x_{18it}^3 - 0,1026378x_{35it}^3 \end{aligned}$$

A partir de estos datos, y sustituyendo los valores de las variables por sus valores en la media, se puede estimar la elasticidad de las variables; elasticidad que se estima globalmente; esto es, para todos los individuos, y para cada individuo. Las estimaciones son las obtenidas en el cuadro que se presenta a continuación.

Variables	Elasticidad en la media	Elasticidad en los establecimientos			
		Hiper	Super de 1000 a 2499 m	Super de 400 a 999 m	Super de 100 a 399 m
Yit-1	0,3966162	0,3966162	0,3966162	0,3966162	0,3966162
x1	-0,21616039	-0,2672565	-0,21487161	-0,21952131	-0,16299209
x2	0,56312984	0,59892547	0,5629051	0,54279606	0,5478922
x9	0,03465849	0,0184701	0,04288532	0,04124237	0,03603615
x13	0,1049743	0,09360696	0,10690042	0,10817127	0,11121868
x14	0,0107228	0,01589194	0,01236608	0,00827449	0,00635869
x17	0,0030935	0,00246308	0,00182556	0,00602082	0,00206454
x18	0,01907337	0,02980321	0,02113396	0,01212244	0,01192138
x28	0,00077987	0,00311947	0	0	0
x35	0,00749487	0,02990723	0	0	0

Tanto el cuadro como el gráfico nos informan de la importancia de la influencia de las variables. De tal modo que por orden de importancia las variables son:

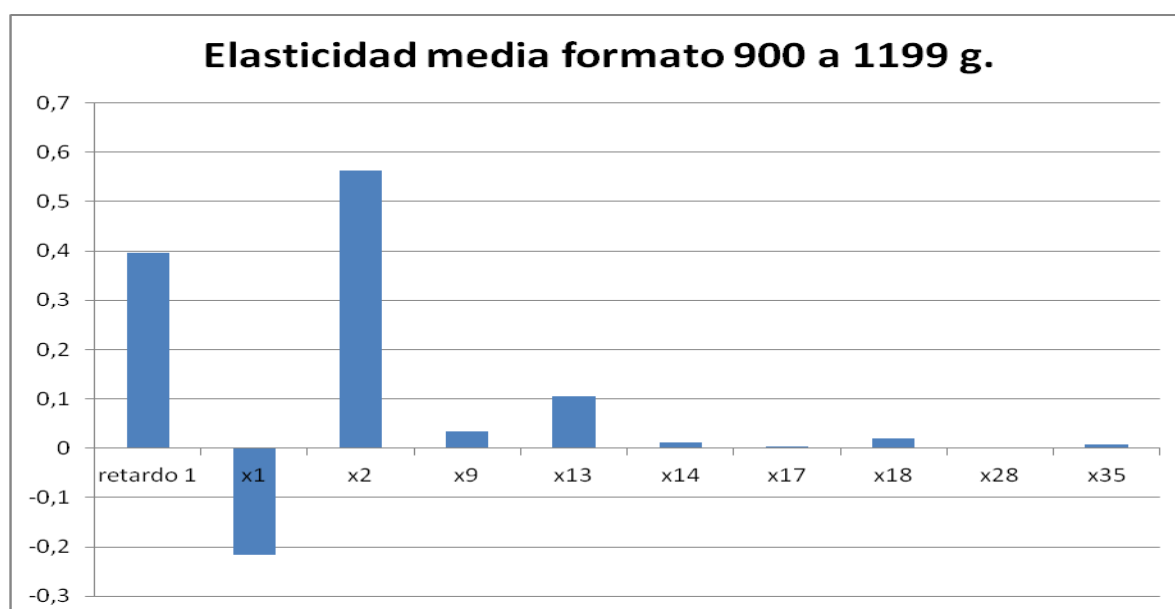
- **Diferencial Precio Principal Competidor: ( $x_2$ )** El signo asociado a la variable es el esperado, positivo. De forma que un diferencial de precio más acusado con el Principal Competidor revertirá en una cuota en tienda más elevada para la Marca 1.
- **Retardo1** Esta variable mide la influencia que tiene la variable a explicar (cuota de mercado) del periodo anterior. El signo asociado es positivo de tal forma que cuotas elevadas en el periodo anterior producen aumentos en la cuota de mercado del periodo actual.
- **Distribución Ponderada Marca 1 ( $x_1$ )** Su signo es negativo y contrario al esperado. La explicación posible puede venir desde la óptica de Marketing. Es posible que aumentos en la distribución ponderada se relacionen con descensos de la cuota en tienda; el motivo, cuando ganamos distribución estamos entrando en tiendas en las que antes la Marca 1 no estaba presente, pero donde sí se vendía el producto; es posible que los consumidores continúen con las marcas que consumían y no cambien a la Marca 1, por lo que su cuota en tienda desciende.



- **Cobertura Reducción Temporal de Precio Marca 1 ( $x_{13}$ )** Esta variable indica en qué grado o con qué intensidad se ha realizado esta promoción de la Marca 1. El signo positivo indica la relación directa entre las acciones comerciales de Reducción Temporal de Precio de la Marca 1 y su cuota en tienda.
- **Cobertura Promoción Folleto y Exposición Especial Marca 1 ( $x_9$ )** La variable mide el grado de participación de los puntos de venta en esta acción emprendida por la Marca 1. Su rango de variación va de 0, cuando ningún establecimiento participa en ella hasta 1, todos los establecimientos participan. El signo es positivo ilustrando la relación directa que existe entre la variable y la cuota en tienda de la Marca 1.
- **Cobertura Promoción sólo Folleto Principal Competidor ( $x_{18}$ )** La variable mide la influencia de un tipo especial de promoción realizada por el principal competidor. Sería lógico esperar un signo negativo, indicando una relación inversa entre las variables; ahora bien, en la estimación aparece signo positivo (relación directa entre las variables); esta circunstancia tiene una interpretación desde el punto de vista de marketing. Una mayor actividad promocional por parte de la competencia llama la atención sobre la categoría de producto (Efecto Llamada), logrando que todas las marcas se beneficien; y es lo que está sucediendo en este caso.
- Las restantes variables tienen un efecto muy parecido sobre la cuota de mercado de la marca 1. Las variables aquí incluidas son:
  - a. **Cobertura Regalo Directo Marca 1 ( $x_{14}$ )**. Valores elevados se asocian a valores elevados en la cuota en tienda de la Marca 1.
  - b. **Cobertura Regalo Directo Marcas Distribución ( $x_{35}$ )** La variable indica que una relación directa entre esta acción realizada por las Marcas de Distribución y la Marca 1. Esta relación de existir, se espera que sea inversa, pero aquí es directa, además, si ya es difícil explicar la razón de esta relación, ahora se vuelve todavía más complicada su interpretación porque esta variable recoge NO LINEALIDADES del

modelo, como lo prueba la existencia como variable significativa de la variable  $x_{35}$  al cubo.

- c. **Cobertura Promoción sólo Exposición Especial Principal Competidor ( $x_{17}$ )**. El signo es positivo, cuando se espera un signo negativo. La explicación, de nuevo desde la óptica del marketing es que este tipo de acción llama la atención sobre la categoría de producto, beneficiando a todas las marcas presentes en el punto de venta entre las que está y de forma destaca la Marca 1.
- d. **Cobertura Regalo Directo Resto De Fabricantes ( $x_{28}$ )** La variable indica la influencia de la promoción de regalo directo realizada por las Marcas de Distribución. Su interpretación es la misma que la variable  $x_{18}$ , incluso en su contradicción aparente.



Cuando analizamos la elasticidad por tipo de establecimientos se aprecian las siguientes características.

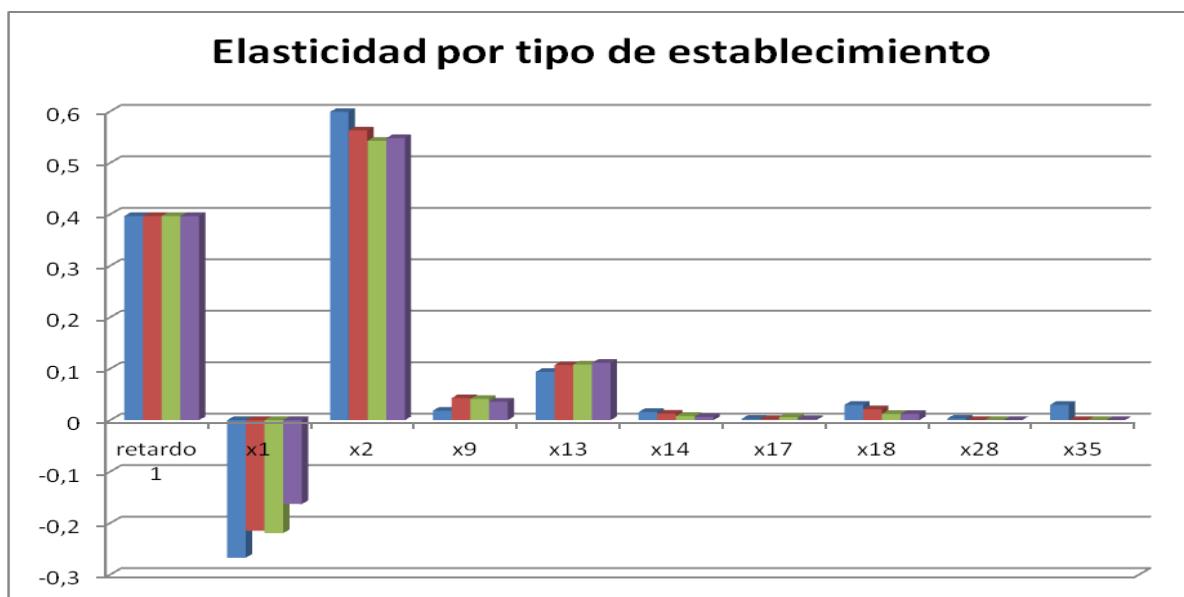
- **Retardo1** Su influencia es la misma para todos los tipos de establecimientos.
- **Diferencial Precio Principal Competidor: ( $x_2$ )** La influencia de esta variable es mayor en los hipermercados que en el resto de establecimientos, ya que es en los hipermercados donde se produce una mayor grado de competencia entre las marcas.
- **Distribución Ponderada Marca 1 ( $x_1$ )** De nuevo es en los hipermercados donde se produce una mayor influencia de esta variable, ya que son los

establecimientos con mayor volumen de venta, y esta variable mide precisamente la calidad de la distribución desde el punto de vista de ventas, estar presente en los establecimientos que más venden produce una distribución ponderada para la marca más elevada que estar presente en otros establecimientos con menor volumen de ventas; por ejemplo los supermercados pequeños.

- **Cobertura Reducción Temporal de Precio Marca 1 ( $x_{13}$ )**. El efecto de este tipo de promoción de la Marca 1 es mayor cuanto más pequeño es el establecimiento. Establecimientos más pequeños muestran un menor grado de competencia, por lo que las acciones comerciales realizadas por la Marca 1 resultan más efectivas que las mismas acciones realizadas en otros establecimientos más grandes.
- **Cobertura Promoción Folleto y Exposición Especial Marca 1 ( $x_9$ )** Esta variable es prácticamente la única que presenta una conducta un tanto diferente a las restantes. Su efecto es mayor en los supermercados grandes y medianos. En buena lógica es donde debería utilizarse con mayor profusión, máxime cuando su coste de utilización es menor que en los hipermercados.
- **Cobertura Promoción sólo Folleto Principal Competidor ( $x_{18}$ )** Esta variable presenta su mayor influencia en los hipermercados; lógico si tenemos en cuenta que es un tipo de acción muy utilizada en estos establecimientos.
- **Cobertura Regalo Directo Marca 1 ( $x_{14}$ )**. De nuevo presenta valores más elevados en los hipermercados porque es allí donde más se utiliza.
- **Cobertura Regalo Directo Marcas Distribución ( $x_{35}$ )** Es en los hipermercados, donde existe una mayor presencia de las marcas de la distribución, donde se produce una mayor influencia de la variable. En los restantes establecimientos la elasticidad es cero, indicando que este tipo de acción no se realiza en los supermercados, bien porque no la hacen o porque no tienen marcas de distribución; es más la primera causa que la segunda, porque poner en marcha una promoción de estas características, lleva aparejada una complejidad logística que no pueden abordar los departamentos de marketing de estos establecimientos.
- **Cobertura Promoción sólo Exposición Especial Principal Competidor ( $x_{17}$ )** Muestra exactamente el mismo comportamiento que variables

anteriores. Mayor presencia en hipermercados que en los demás establecimientos y su influencia es directa al tamaño del punto de venta, cuanto más grande es el establecimiento más influye esta variable.

- **Cobertura Regalo Directo Resto De Fabricantes ( $x_{28}$ )** En este caso la variable muestra valores distintos de cero sólo en los hipermercados. En los restantes establecimientos es cero, bien porque no existe la presencia de Resto de Fabricantes o porque éstos no utilizan este tipo de promoción.



### 7.1.5.- Elasticidad en el modelo de 1200 A 1499 g.

$$Lny_{it} = \delta + \theta_4 Lny_{it-4} + \beta_9 x_{9it} + \beta_{13} x_{13it} + \beta_{17} x_{17it} + \theta_{201} Lny_{it-1}^2 + \beta_{216} x_{16it}^2 + \theta_{304} Lny_{it-4}^3 + \beta_{313} x_{13it}^3 + \beta_{316} x_{16it}^3 + u_{it}$$

#### Efectos marginales

$$\frac{\partial Lny_{it}}{\partial x_{9it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{9it}}{x_{9it}}} = \beta_9 \quad ; \quad \frac{\partial y_{it}}{\partial x_{9it}} = \beta_9 y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{13it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{13it}}{x_{13it}}} = \beta_{13} + 3\beta_{313} x_{13it}^2 \quad ; \quad \frac{\partial y_{it}}{\partial x_{13it}} = (\beta_{13} + 3\beta_{313} x_{13it}^2) y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{16it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{16it}}{x_{16it}}} = 3\beta_{316} x_{16it}^2 + 2\beta_{216} x_{16it} \quad ; \quad \frac{\partial y_{it}}{\partial x_{16it}} = (3\beta_{316} x_{16it}^2 + 2\beta_{216} x_{16it}) y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{17it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{17it}}{x_{17it}}} = \beta_{17} \quad ; \quad \frac{\partial y_{it}}{\partial x_{17it}} = \beta_{17} y_{it}$$

$$\frac{\partial Lny_{it}}{\partial Lny_{it-1}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial y_{it-1}}{y_{it-1}}} = 2\theta_{201} \frac{y_{it-1}}{y_{it-1}^2} \quad ; \quad \frac{\partial y_{it}}{\partial y_{it-1}} = \left( 2\theta_{201} \frac{y_{it-1}}{y_{it-1}^2} \right) \frac{y_{it}}{y_{it-1}}$$

$$\frac{\partial Lny_{it}}{\partial Lny_{it-4}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial y_{it-4}}{y_{it-4}}} = \theta_4 + 3\theta_{304} \frac{y_{it-4}^2}{y_{it-4}^3} \quad ; \quad \frac{\partial y_{it}}{\partial y_{it-4}} = \left( \theta_4 + 3\theta_{304} \frac{y_{it-4}^2}{y_{it-4}^3} \right) \frac{y_{it}}{y_{it-4}}$$

#### Elasticidad

$$\varepsilon_{x_{9it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{9it}}{x_{9it}}} \frac{x_{9it}}{y_{it}} \quad ; \quad \varepsilon_{x_{9it}} = \beta_9 y_{it} \frac{x_{9it}}{y_{it}} \quad ; \quad \varepsilon_{x_{9it}} = \beta_9 x_{9it}$$

$$\varepsilon_{x_{13it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{13it}}{x_{13it}}} \frac{x_{13it}}{y_{it}} \quad ; \quad \varepsilon_{x_{13it}} = (\beta_{13} + 3\beta_{313} x_{13it}^2) y_{it} \frac{x_{13it}}{y_{it}} \quad ; \quad \varepsilon_{x_{13it}} = \beta_{13} x_{13it} + 3\beta_{313} x_{13it}^3$$

$$\varepsilon_{x_{16it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{16it}}{x_{16it}}} \frac{x_{16it}}{y_{it}} \quad ; \quad \varepsilon_{x_{16it}} = (3\beta_{316} x_{16it}^2 + 2\beta_{216} x_{16it}) y_{it} \frac{x_{16it}}{y_{it}} \quad ; \quad \varepsilon_{x_{16it}} = 3\beta_{316} x_{16it}^3 + 2\beta_{216} x_{16it}^2$$

$$\varepsilon_{x_{17it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{17it}}{x_{17it}}} \frac{x_{17it}}{y_{it}} \quad ; \quad \varepsilon_{x_{17it}} = \beta_{17} y_{it} \frac{x_{17it}}{y_{it}} \quad ; \quad \varepsilon_{x_{17it}} = \beta_{17} x_{17it}$$

$$\varepsilon_{y_{it-1}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial y_{it-1}}{y_{it-1}}} \frac{y_{it-1}}{y_{it}} \quad ; \quad \varepsilon_{y_{it-1}} = \left( 2\theta_{201} \frac{y_{it-1}}{y_{it-1}^2} \right) \frac{y_{it}}{y_{it-1}} \frac{y_{it-1}}{y_{it}} \quad ; \quad \varepsilon_{y_{it-1}} = 2\theta_{201} \frac{1}{y_{it-1}}$$

$$\varepsilon_{y_{it-4}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial y_{it-4}}{y_{it-4}}} \frac{y_{it-4}}{y_{it}} \quad ; \quad \varepsilon_{y_{it-4}} = \left( \theta_4 + 3\theta_{304} \frac{y_{it-4}^2}{y_{it-4}^3} \right) \frac{y_{it}}{y_{it-4}} \frac{y_{it-4}}{y_{it}} \quad ; \quad \varepsilon_{y_{it-4}} = \theta_4 + \frac{3\theta_{304}}{y_{it-4}}$$

$$\begin{aligned} \text{Ln}\hat{y}_{it} = & -0,9038911 + 0,3815741\text{Ln}y_{it-4} + 0,0142113x_{9it} + 0,0050983x_{13it} + 0,0027367x_{17it} \\ & -0,1312634\text{Ln}y_{it-1}^2 - 0,0012601x_{16it}^2 - 0,010813\text{Ln}y_{it-4}^3 + 0,00000117x_{13it}^3 + 0,0000582x_{16it}^3 \end{aligned}$$

La estimación del modelo nos permite hacer inferencia sobre la elasticidad en la media de las variables y en la media para cada tipo de establecimiento. Ambas estimaciones están realizadas en el cuadro que viene a continuación.

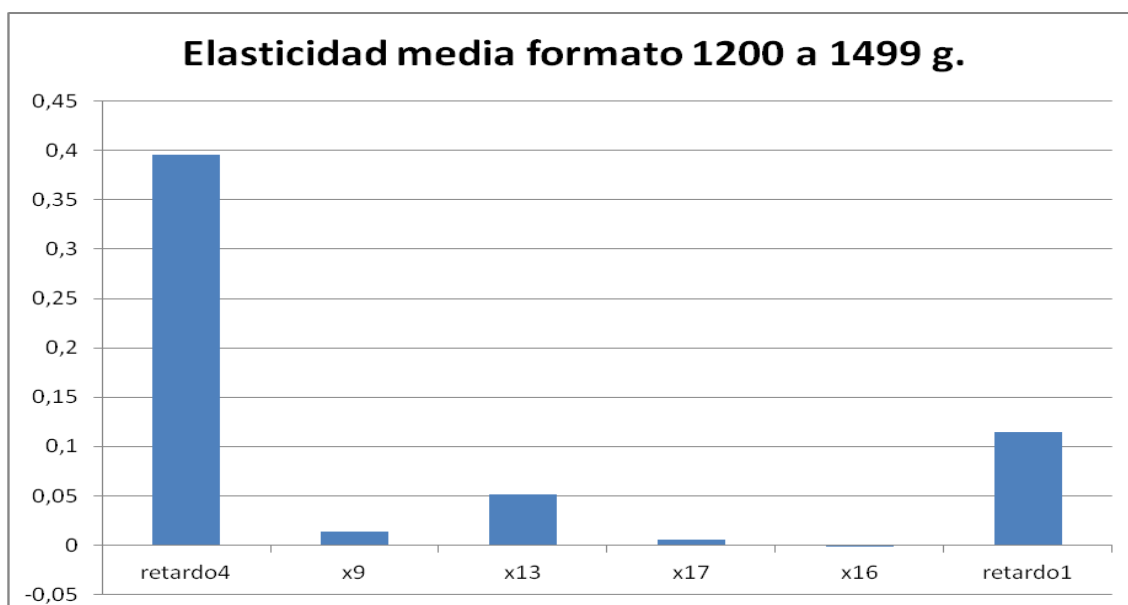
Variables	Elasticidad en la media	Elasticidad en los establecimientos			
		Hiper	Super de 1000 a 2499 m	Super de 400 a 999 m	Super de 100 a 399 m
<b>Yit-4</b>	0,3956788	0,39245637	0,39425711	0,39760116	0,4013915
<b>x9</b>	0,01392721	0,00662534	0,01441314	0,02341526	0,01125509
<b>x13</b>	0,05133169	0,03650091	0,04901314	0,0476045	0,07427614
<b>x17</b>	0,00569589	0,00284808	0,01176599	0,00615496	0,00201455
<b>x16</b>	-0,0001166	-0,00178023	0	0	0
<b>Yit-1</b>	0,11502271	0,08788123	0,10281923	0,13160934	0,16467166

Tal y como se aprecia tanto en el cuadro como en el gráfico, la importancia de las variables respecto a su efecto sobre la cuota de mercado de la marca 1 es:

- **Retardo4** Mide la influencia de la variable cuota de mercado de 4 periodos atrás. Su signo asociado es positivo estableciéndose una relación directa entre ambas variables; aumentos en una suponen aumentos en la otra y viceversa. Por tanto el trabajo bien realizado en el pasado tiene un efecto positivo sobre el presente, aunque si bien es cierto que hay que esperar cuatro periodos para comprobarlo.
- **Retardo1 al cuadrado.** Mide la influencia que tiene la variable cuota de mercado del periodo anterior, pero recogiendo un problema de no linealidad presente en el modelo. El signo negativo asociado nos hace pensar en la existencia de un máximo en la función.
- **Retardo4 al cubo.** Recoge la presencia de no linealidades en el modelo.
- **Cobertura Reducción Temporal De Precio Marca 1 (x13)** Esta variable mide el efecto positivo que supone realizar este tipo de promoción. Es de

resaltar que lo que la variable mide no es la cuantía de la promoción, sino que la promoción sea muy visible; que esté en muchos establecimientos a la vez.

- **Cobertura Promoción Folleto y Exposición Especial Marca 1 ( $x_9$ )** El signo positivo es el que efectivamente tiene que ser. Valores bajos de la variable (gran cantidad de establecimientos que realizan este tipo de acción) se asocian a valores bajos de la cuota en tienda y valores elevados se asocian a cuota en tienda elevada de la Marca 1; por lo que ambas variables varían en el mismo sentido.
- **Cobertura Promoción sólo Exposición Especial Principal Competidor ( $x_{17}$ )**. La relación entre ambas variables es negativa, cuando el signo esperado es positivo. Nuevamente la justificación puede venir desde el punto de vista del marketing. Esta acción realizada por la competencia, llama la atención sobre la categoría de producto, movilizand o a los compradores, que materializan sus opciones de compra no exclusivamente en la marca que realiza la acción, sino que otras marcas, como la Marca 1, se ven beneficiadas directamente por esta acción.
- **Cobertura Promoción Folleto y Exposición Especial Principal Competidor al Cuadrado y al cubo ( $x_{16}$ )**. Recoge la influencia de no linealidades presentes en el modelo estimado.

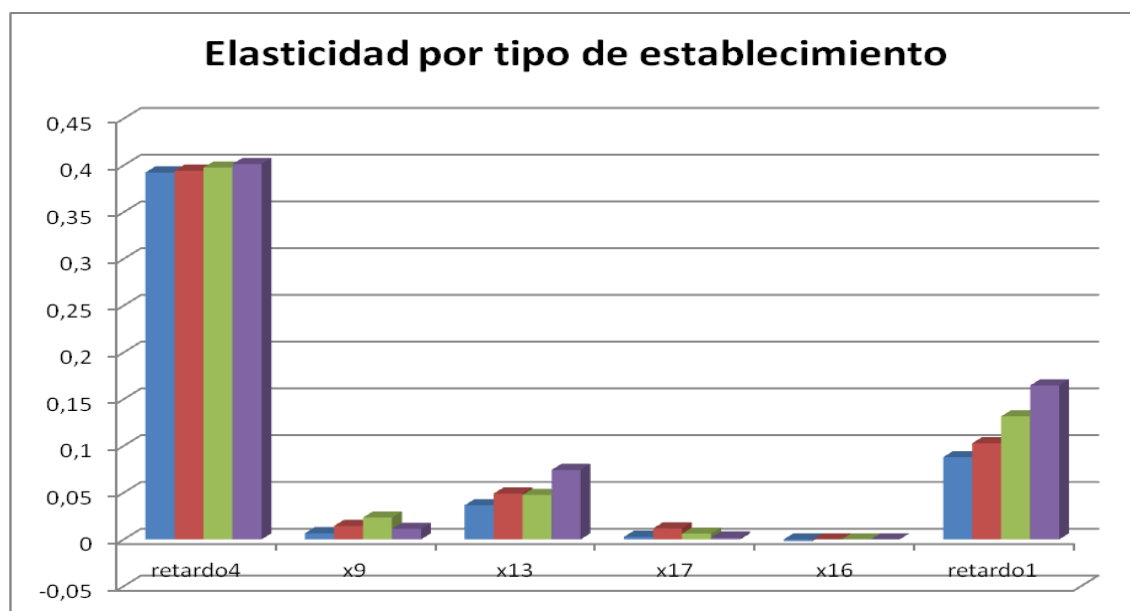


Cuando analizamos la elasticidad por tipos de establecimientos, las conclusiones que hemos ido obteniendo se van matizando, pudiendo comprobarse que los efectos son distintos según de que establecimiento estemos hablando.

Así puede comprobarse que:

- Las elasticidades de las variables **retardo 4** y **retardo 1 al cuadrado** no son constantes, cuando lo habitual en otros formatos ha sido precisamente eso; y además su influencia es mayor cuanto más pequeño es el establecimiento. Parece por tanto que los establecimientos más pequeños se comportan de una forma más fiel que los grandes, seguramente influenciado por el grado de competencia existente en ellos, ya que en los pequeños la competencia es menor y la influencia de las acciones realizadas en el pasado es mayor.
- La elasticidad de la variable **Cobertura Reducción Temporal de Precio Marca 1 ( $X_{13}$ )** es mayor en los establecimientos más pequeños que en los más grandes.
- La elasticidad de la variable **Cobertura Promoción Folleto y Exposición Especial Marca 1 ( $x_9$ )** tiene especial importancia en los establecimientos medianos. Es en estos establecimientos donde se dan la circunstancia de menor clima competitivo que hace que la variable tenga un mayor efecto sobre la cuota en tienda de la Marca 1.
- Elasticidad de la variable **Cobertura Promoción sólo Exposición Especial Principal Competidor ( $X_{17}$ )**. Es en los establecimientos medianos donde esta variable adquiere más importancia, hablándonos de un clima competitivo más benigno que aumenta los rendimientos de esta acción de marketing en estos tipos de establecimientos.
- Elasticidad de la variable **Cobertura Promoción Folleto y Exposición Especial Principal Competidor al cuadrado y al cubo ( $x_{16}$ )** Esta variable sólo tiene efecto en los hipermercados, debido quizás a al mayor grado de competencia existente en estos establecimientos.





### 7.1.6.- Elasticidad en el modelo de Hipermercados

$$Lny_{it} = \delta + \theta_1 Lny_{it-1} + \theta_4 Lny_{it-4} + \beta_9 x_{9it} + \beta_{10} x_{10it} + \beta_{11} x_{11it} + \beta_{24} x_{24it} + \beta_{26} x_{26it} + \beta_{27} x_{27it} + \beta_{32} x_{32it} + \beta_{33} x_{33it} + u_{it}$$

#### Efectos marginales

$$\frac{\partial Lny_{it}}{\partial x_{9it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{9it}}{x_{9it}}} = \beta_9 ; \frac{\partial y_{it}}{\partial x_{9it}} = \beta_9 y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{10it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{10it}}{x_{10it}}} = \beta_{10} ; \frac{\partial y_{it}}{\partial x_{10it}} = \beta_{10} y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{11it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{11it}}{x_{11it}}} = \beta_{11} ; \frac{\partial y_{it}}{\partial x_{11it}} = \beta_{11} y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{24it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{24it}}{x_{24it}}} = \beta_{24} ; \frac{\partial y_{it}}{\partial x_{24it}} = \beta_{24} y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{26it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{26it}}{x_{26it}}} = \beta_{26} ; \frac{\partial y_{it}}{\partial x_{26it}} = \beta_{26} y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{27it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{27it}}{x_{27it}}} = \beta_{27} ; \frac{\partial y_{it}}{\partial x_{27it}} = \beta_{27} y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{32it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{32it}}{x_{32it}}} = \beta_{32} ; \frac{\partial y_{it}}{\partial x_{32it}} = \beta_{32} y_{it}$$

$$\frac{\partial Lny_{it}}{\partial x_{33it}} = \frac{\frac{\partial y_{it}}{y_{it}}}{\frac{\partial x_{33it}}{x_{33it}}} = \beta_{33} ; \frac{\partial y_{it}}{\partial x_{33it}} = \beta_{33} y_{it}$$

$$\frac{\partial Lny_{it}}{\partial Lny_{it-1}} = \frac{\frac{\frac{\partial y_{it}}{y_{it}}}{y_{it-1}}}{\frac{\partial y_{it-1}}{y_{it-1}}} = \theta_1 ; \frac{\partial y_{it}}{\partial y_{it-1}} = \theta_1 \frac{y_{it}}{y_{it-1}}$$

$$\frac{\partial Lny_{it}}{\partial Lny_{it-4}} = \frac{\frac{\frac{\partial y_{it}}{y_{it}}}{y_{it-4}}}{\frac{\partial y_{it-4}}{y_{it-4}}} = \theta_4 ; \frac{\partial y_{it}}{\partial y_{it-4}} = \theta_4 \frac{y_{it}}{y_{it-4}}$$

## Elasticidad

$$\begin{aligned}
\varepsilon_{x_{9it}} &= \frac{\partial y_{it}}{\partial x_{9it}} \frac{x_{9it}}{y_{it}} ; \varepsilon_{x_{9it}} = \beta_9 y_{it} \frac{x_{9it}}{y_{it}} ; \varepsilon_{x_{9it}} = \beta_9 x_{9it} \\
\varepsilon_{x_{10it}} &= \frac{\partial y_{it}}{\partial x_{10it}} \frac{x_{10it}}{y_{it}} ; \varepsilon_{x_{10it}} = \beta_{10} y_{it} \frac{x_{10it}}{y_{it}} ; \varepsilon_{x_{10it}} = \beta_{10} x_{10it} \\
\varepsilon_{x_{11it}} &= \frac{\partial y_{it}}{\partial x_{11it}} \frac{x_{11it}}{y_{it}} ; \varepsilon_{x_{11it}} = \beta_{11} y_{it} \frac{x_{11it}}{y_{it}} ; \varepsilon_{x_{11it}} = \beta_{11} x_{11it} \\
\varepsilon_{x_{24it}} &= \frac{\partial y_{it}}{\partial x_{24it}} \frac{x_{24it}}{y_{it}} ; \varepsilon_{x_{24it}} = \beta_{24} y_{it} \frac{x_{24it}}{y_{it}} ; \varepsilon_{x_{24it}} = \beta_{24} x_{24it} \\
\varepsilon_{x_{26it}} &= \frac{\partial y_{it}}{\partial x_{26it}} \frac{x_{26it}}{y_{it}} ; \varepsilon_{x_{26it}} = \beta_{26} y_{it} \frac{x_{26it}}{y_{it}} ; \varepsilon_{x_{26it}} = \beta_{26} x_{26it} \\
\varepsilon_{x_{27it}} &= \frac{\partial y_{it}}{\partial x_{27it}} \frac{x_{27it}}{y_{it}} ; \varepsilon_{x_{27it}} = \beta_{27} y_{it} \frac{x_{27it}}{y_{it}} ; \varepsilon_{x_{27it}} = \beta_{27} x_{27it} \\
\varepsilon_{x_{32it}} &= \frac{\partial y_{it}}{\partial x_{32it}} \frac{x_{32it}}{y_{it}} ; \varepsilon_{x_{32it}} = \beta_{32} y_{it} \frac{x_{32it}}{y_{it}} ; \varepsilon_{x_{32it}} = \beta_{32} x_{32it} \\
\varepsilon_{x_{33it}} &= \frac{\partial y_{it}}{\partial x_{33it}} \frac{x_{33it}}{y_{it}} ; \varepsilon_{x_{33it}} = \beta_{33} y_{it} \frac{x_{33it}}{y_{it}} ; \varepsilon_{x_{33it}} = \beta_{33} x_{33it} \\
\varepsilon_{y_{it-1}} &= \frac{\partial y_{it}}{\partial y_{it-1}} \frac{y_{it-1}}{y_{it}} ; \varepsilon_{y_{it-1}} = \theta_1 \frac{y_{it}}{y_{it-1}} \frac{y_{it-1}}{y_{it}} ; \varepsilon_{y_{it-1}} = \theta_1 \\
\varepsilon_{y_{it-4}} &= \frac{\partial y_{it}}{\partial y_{it-4}} \frac{y_{it-4}}{y_{it}} ; \varepsilon_{y_{it-4}} = \theta_4 \frac{y_{it}}{y_{it-4}} \frac{y_{it-4}}{y_{it}} ; \varepsilon_{y_{it-4}} = \theta_4
\end{aligned}$$

Este es el único modelo donde la elasticidad estimada, en lugar de calcularse para los establecimientos donde se vende el producto, se calcula para los diferentes formatos que están presentes en los hipermercados.

La estructura estimada de este modelo es la que se muestra a continuación:

$$\begin{aligned}
Ln\hat{y}_{it} = & -1,023057 + 0,4981634Ln y_{it-1} + 0,1569713Ln y_{it-4} + 0,006346x_{9it} + 0,0072708x_{10it} \\
& + 0,0044697x_{11it} - 0,0027524x_{24it} - 0,0010845x_{26it} - 0,0011324x_{27it} - 0,001221x_{32it} - 0,0015896x_{33it}
\end{aligned}$$

Básicamente hablamos de un modelo que desde el punto de vista estadístico no presenta diferencias respecto a los anteriores. La diferencia reside en que los individuos que forman el panel, no son establecimientos, sino los formatos presentes en los hipermercados.

Las elasticidades estimadas en la media globalmente, para todos los formatos de la marca 1, están en la primera columna del gráfico siguiente. Esta información se complementa con las estimaciones de las elasticidades para los distintos formatos; este análisis se realiza para 6 formatos diferentes.

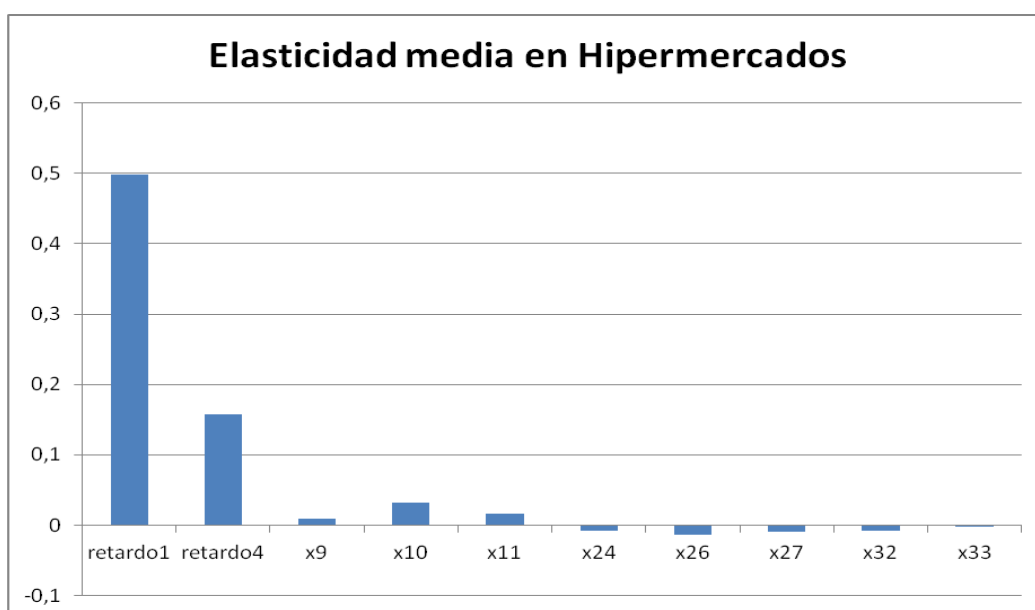
	Elasticidad en la media	Elasticidad en los formatos					
		Hasta 350	De 350 a 499	De 500 a 899	De 900 a 1199	De 1200 a 1499	De 3000 a 4999
Yit-1	0,4981634	0,4981634	0,4981634	0,4981634	0,4981634	0,4981634	0,4981634
Yit-4	0,1569713	0,1569713	0,1569713	0,1569713	0,1569713	0,1569713	0,1569713
x9	0,00860875	0	0	0,00375268	0,01210892	0,00295852	0,03283237
x10	0,03190335	0,00096376	0	0,00541745	0,01761055	0,00590737	0,16152097
x11	0,01625047	0	0	0,00449836	0,0471961	0,01110283	0,03470555
x24	-0,00748186	-0,00552755	-0,01002638	0	-0,00414976	0	-0,02518747
x26	-0,01294943	-0,00261452	-0,00468499	-0,00285964	0	0	-0,06753743
x27	-0,00995646	-0,0118081	-0,01902956	-0,01840381	-0,00624497	0	-0,00425231
x32	-0,00739687	0	-0,00215858	-0,00250219	-0,02224618	-0,00720789	-0,0102664
x33	-0,00032636	0	-0,00195816	0	0	0	0

Del análisis del cuadro y gráfico se obtienen las conclusiones siguientes:

- Las variables que muestran una elasticidad más importante son **Retardo1** y **Retardo4**. Indicando en ambos casos que cuotas de mercado, y por tanto ventas, elevadas en el pasado producen efectos positivos en la cuota de mercado actual; ahora bien, ese efecto sólo se produce con un máximo de 4 retardos, por lo tanto la memoria comercial de la marca parece bastante finita.
- La siguiente elasticidad más importante es la referida a la variable **Cobertura Promoción sólo Exposición Especial Marca 1 (X<sub>10</sub>)** que mide el efecto sobre la cuota en tienda de la Marca 1 de este tipo de promoción. El signo de la elasticidad es positivo, esto es, existe una relación positiva y directa entre la variable y la cuota en tienda de la Marca 1.
- En cuarto lugar aparece la elasticidad debida a la **Cobertura Promoción Solo Folleto Marca 1 (X<sub>11</sub>)**.
- La elasticidad debida a la **Cobertura Promoción Folleto y Exposición Especial Marca 1 (X<sub>9</sub>)** aparece en quinta posición. Son este tipo de acciones realizadas individualmente o en conjunto las que producen un efecto positivo sobre la marca en los hipermercados.
- En sexto lugar, por orden de importancia, aparece la elasticidad de la variable **Cobertura Otras Promociones Resto de Fabricantes (X<sub>26</sub>)**. Con esta

elasticidad medimos la relación entre el grupo de otras promociones realizadas por la competencia encarnada en el resto de fabricantes y la cuota en tienda de la Marca 1. El signo es NEGATIVO, lógico, ya que una mayor actividad promocional de la competencia esta relacionada con una menor cuota de mercado de la Marca 1. La importancia competitiva que juegan el resto de fabricantes vuelve a ponerse de manifiesto en esta variable.

- La séptima posición por importancia de elasticidad está ocupada por la variable ***Cobertura Reduccion Temporal de Precios Resto De Fabricantes (X<sub>27</sub>)***. Esta variable mide una relación muy similar a la anterior; la diferencia está en el tipo de promoción considerado. Su signo NEGATIVO nos indica de la relación inversa entre las acciones del resto de fabricantes y la cuota en tienda de la Marca 1.
- La posición siguiente, en valor absoluto, aunque con signo negativo está ocupada por la elasticidad de la variable ***Cobertura Promoción sólo Exposición Especial Resto De Fabricantes (X<sub>24</sub>)***. Relación inversa entre este tipo de promoción realizada por la competencia y la cuota en tienda de la Marca 1, de manera que una mayor actividad promocional del resto de fabricantes afectará negativamente a la cuota en tienda de la marca estudiada.
- La penúltima posición según su importancia está ocupada por la elasticidad de la variable ***Cobertura Promoción Solo Folleto Marcas Distribución (X<sub>32</sub>)***. En esta oportunidad se mide la intensidad de la relación entre la intensidad promocional, para el tipo “SÓLO FOLLETO”, realizada por las marcas de la distribución y la Marca 1. Esta relación es inversa, de tal modo que a mayor actividad promocional de las marcas de distribución, menor es la cuota en tienda de la Marca 1.
- En última posición aparece la elasticidad de la variable ***Cobertura Otras Promociones Marcas Distribución (X<sub>33</sub>)***. Nuevamente tenemos una relación inversa entre la acciones realizadas por las marcas de la distribución y la cuota en tienda de la Marca 1.

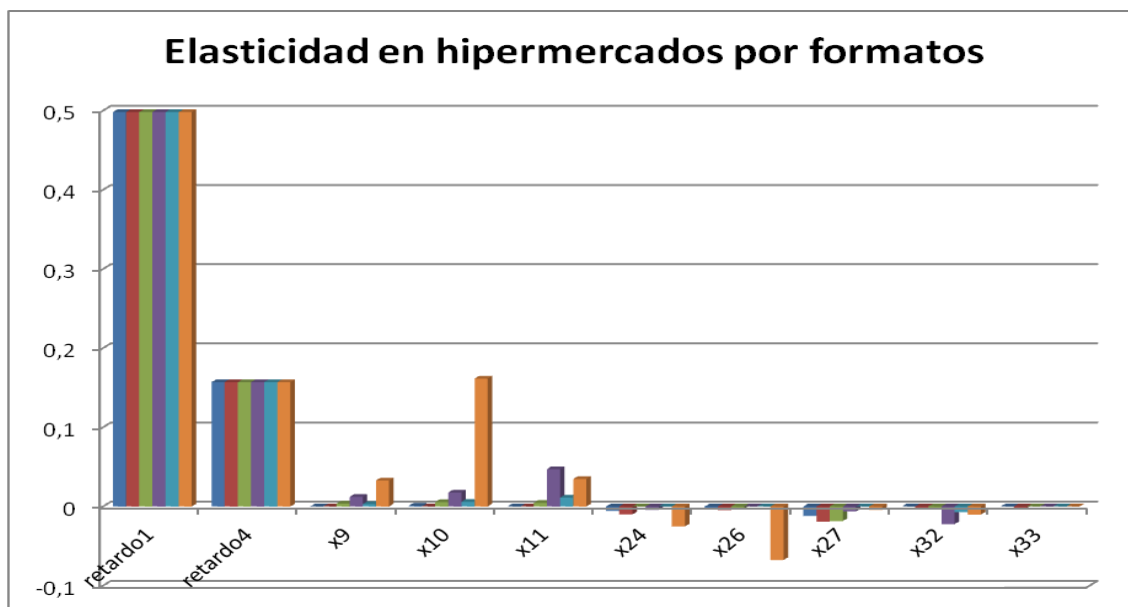


Cuando analizamos la elasticidad en función de los formatos se observan diferencias apreciables en su intensidad, aunque el signo es el mismo al del valor medio de la elasticidad estimada.

- Así, mientras en las variables retardo 1 y retardo 4 la elasticidad es constante no sucede lo mismo con las demás variables.
- La elasticidad de la variable **Cobertura Promoción Solo Exposición Especial Marca 1 ( $X_{10}$ )** alcanza los valores más elevados en el formato de 3000 a 4999 g.; y esto mismo sucede con las elasticidades de las variables **Cobertura Promoción Folleto y Exposición Especial Marca 1 ( $X_9$ )**, **Cobertura Promoción sólo Exposición Especial Resto de fabricantes ( $X_{24}$ )** y **Cobertura Otras Promociones Resto de Fabricantes ( $X_{26}$ )**, demostrando que es en el formato de mayor tamaño donde se concentra la mayor competencia comercial entre las marcas en los hipermercados.
- En el formato de 500 a 899 g. las variables que muestran una mayor incidencia son **Cobertura Reducción Temporal de Precios Resto De Fabricantes ( $X_{27}$ )** y **Cobertura Promoción Solo Folleto Marcas Distribución ( $X_{32}$ )**. Esta concentración de acciones vuelve a poner de

manifiesto que este formato es el segundo por grado de competencia comercial.

- Las elasticidades de las variables **Cobertura Promoción sólo Folleto Marca 1 ( $X_{11}$ )** y **Cobertura Otras Promociones Marcas Distribución ( $X_{33}$ )** tienen especial relevancia en los formatos de 900 a 1199 g. , la primera y en el formato de 350 a 499 g. , exclusivamente, la segunda.



- Es sintomático que el formato de Hasta 350 g. no aparezca en primera posición en ninguna de las elasticidades estimadas. Esta circunstancia nos hace pensar que este formato está un poco al margen de las luchas competitivas entre las marcas. Esta lucha se centra en los formatos de mayor tamaño, que son los que se reparten el mayor volumen de ventas.

## 8.- BIBLIOGRAFÍA

- Adkins, L.E. C. ; Carter Hill, R. (2011) *Using Stata por Principles of Econometrics* 4<sup>o</sup> edition Wiley
- Aaker, D.A. (1975) "An Advertising Decision Model" *Journal of Marketing* 12(1) 37-45
- Abraham, M.M., Lodish, L.M. (1989) "PROMOTIONSCAN: A system for improving promotion productivity for retaililrs and system for improving promotion productivity for retailers and manufacturers using scanner store and household panel data" *Working Paper #89,007*,Marketing Departament, The Wharton School of the University of Pennsylvania, Philadelphia, PA:
- Arellano, M (1987) "Computing Robust Standard Errors for Within-Groups Estimators" *Oxford Bulletin of Economics and Statistics* 49 431-434
- Asmus, G., Farley, J.U., Lehmann, D.R.(1984) "How Advertising Affects Sales: Meta-Analysis of Econometric Results" *Journal of Marketing Research* 21(1) 65-74
- Baltagi, B.H. (2001)*Econometric Analysis of Panel Data* 2<sup>o</sup> ed. John Wiley and Sons, LTD
- Baltagi, B.H., Li, Q. (1995) "Testing AR(1) against MA(1) disturbances in an error component model" *Journal of Econometrics*, 68 133-151
- Baltagi, B.H., Li, Q. (1991) "A transformation that will circumvent the problem of autocorrelation in an error-component model" *Journal of Econometrics*, 48 385–393
- Baltagi,B.H., Li, Q. (1997) "Monte Carlo Results on Pure and Pretest Estimators of an Error Component Model with Autocorrelated Disturbances" *ANNALES ECONOMIE ET DE STATISTIQUE* 48 69-82
- Bass, F.M. (1969) "A new product growth model for consumer durables" *Management Science* 15 215-227
- Bass, F.M., R.D. Buzzel, M.R. Greene et al. Eds. (1961) *Mathematical Models and Methods in Marketing*, Homewood, Irwin, IL.
- Bera, A.K., Sosa-Escudero,W., Yoon, M. (2001) "Tests for the error component model in the presence of local misspecification" *Journal of Econometrics*, 101, 1-23
- Berenblut, I.I. and Webb, G.I. (1973) "A New Test for Autocorrelated Errors in the Linear Regression Model" *Journal of the Royal Statistical Society. Series B* (Methodological), 35 33-50
- Bhargava,A., Franzini, L. y Narendranathan, W. (1982) "Serial correlation and the fixed effects model" *Review of Economic Studies*, 49 533-549



Bijmolt, T.H.A., van Heerde, H.J., Pieters, R.G.M. (2005) "New Empirical Generalizations on the Determinants of Price Elasticity" *Journal of Marketing Research*, 42 May 141-156

Blattberg, R.C., Glazer, R., Little, J.D.C. Eds. (1994) *The Marketing Information Revolution* Harvard Business School Press, Boston, MA.

Boletín Económico del ICE (2011) "Capítulo Análisis por formatos y subsectores" *Boletín Económico del ICE* N° 3015, agosto 99-126

Burke, S. P., Godfrey, L. G., and Tremayne, A. R. (1990) "Testing AR(1) Against MA(1) Disturbances in the Linear Regression Model: An Alternative Procedure" *The Review of Economic Studies* 135-145

Buzzell, R.D. (1964) *Mathematical Models and Marketing Management* Harvard University, Division of Research, Boston, MA

Carter Hill, R; Griffiths, W. E.; Lim, G.C (2012) *Principles of Econometrics* Wiley 4<sup>o</sup> ed.

Chow, G. C. (1960) "Tests of Equality Between Sets of Coefficients in Two Linear Regressions" *Econometrica*, 28 591-605

Christen, M., Gupta, S., Porter, J.C., Staelin, R., Wittink, D.R. (1997) "Using Market Level Data to Understand Promotion Effects in a Nonlinear Model" *Journal of Marketing Research*, 34 p 322-334

Clarke, D.G. (1976) "Econometric Measurement of the Duration of Advertising Effects on Sales" *Journal of Marketing Research* 18 345-357

Dorfman, R., Steiner, P.O., (1954) "Optimal advertising and optimal quality" *The American Economic Review* 44 826-836

Frank, J.F., A.A. Kuehn, W.F. Massy, Eds. (1962) *Quantitative Techniques in Marketing Analysis* Homewood, Irwin, IL.

Fuller, W.A., Battese, G.E. (1974) "Estimation of linear models with crossed-error structure" *Journal of Econometrics*, 2 67-78

Gujarati, D; Porter, D (2010) *Econometría* Mc Graw Hill 5<sup>o</sup> ed.

Hausman, J.A. (1978) "Specification Test in Econometrics" *Econometrica*, 46 1251-1271

Hollis, N. (1995) "Like it or not, liking is not enough" *Journal of Advertising Research*. Vol. 35 N° 5, September/October jul-16

Hsiao, C (2002) *Analysis of Panel Data* Cambridge University Press

- Im, K. S., S. C. Ahn, P. Schmidt y J.M. Wooldridge (1999) "Efficient Estimation of Panel Data Models with Strictly Exogenous Explanatory Variables" *Journal of Econometrics*, 93 177 - 201
- Kiefer, N. M. (1980) "Estimation of Fixed Effect Models for Time Series of Cross-Section with Arbitrary Intertemporal Covariance" *Journal of Econometrics*, 14 195 - 202
- King, M.L. (1983) "Testing for autoregressive against moving average errors in the linear regression model" *Journal of Econometrics*, 21 35-51
- Kiviet, J.K (1995) "On bias, inconsistency, and efficiency of various estimators in dynamic panel data models" *Journal of Econometrics* 68 53-78
- Kmenta, J. (1986) *Elements of Econometrics* Macmillan
- Kotler, P.H. (1971) *Marketing Decisions Making: A Model Building Approach* Holt, Rinehart and Winston, New York
- Kotler, P.H. (1965) "Competitive Strategies for New Product Marketing Over the Life Cycle" *Management Science* 12 104-119
- Lee C. Adkins, R. Carter Hill (2011) *Using Stata por Principles of Econometrics 4<sup>o</sup> edition*Wiley 4<sup>o</sup> ed.
- Leeflang, P. S. H (2008) "Modeling Competitive Reaction Effects" *Schmalenbach Business Review* (SBR). Oct2008, Vol. 60 Issue 4 322-358
- Leeflang, P. S. H., Wittink, D.R. (1992) "Diagnosing competitive reactions using (aggregated) scanner data" *International Journal of Research in Marketing* Mar92, Vol. 9 Issue 1, 39-57
- Leeflang, P. S. H., Wittink, D.R. (1996) "Competitive Reaction Versus Consumer Response: Do Managers Overreact?" *International Journal of Research in Marketing*, 13 103-119
- Leeflang, P.S.H. (1974) *Mathematical Models in Marketing, a Survey, the Stage of Development, Some Extensions and Applications* H.E. Stenfert Kroese, Leiden
- Leeflang, P.S.H., Wittink, D.R. (2000) "Building models for marketing decisions: Past, present and future" *International Journal of Research in Marketing* 17 105-126
- Leone, R.P., Schultz, R.L. (1980) "A Study of Marketing Generalizations" *Journal of Marketing Research* 44(Winter) oct-18
- Levene, H. (1960) *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, NI. Olkin et al. eds. Stanford University Press 278-292

- Lilien, G.L., Kotler, P.H., Moorthy, K.S. (1992) *Marketing Models* Prentice Hall, Englewood Cliffs, NJ.
- Little, J.D.C. (1979 b) "Decision Support Systems for Marketing Managers" *Journal of Marketing* 43(3) sep-26
- Little, J.D.C. (1979) "Aggregate Advertising Models: the State of the Art" *Operations Research* 27(4) July-August 629-667
- Little, J.D.C. (1970) "Models and Managers: The Concept of a Decisions Calculus" *Management Science*, 16 B466-B485
- Lodish, L.M. (1971) CALLPLAN: "An Interactive Salesman's Call Planning System" *Management Science* 18 (4 Par II) 25-40
- Massy, W.F., Montgomery, D.B., Morrison, D.G. (1970) *Stochastic Models of Buying Behavior* M.I.T. Press, Boston, MA.
- Mulhern, F.J., Leone, R.P. (1991) "Implicit price bundling of retail products: a multiproduct approach to maximizing store profitability" *The Journal of Marketing* 55, October 63-76
- Nagar, A. L. (1959) "The Bias and Moment Matrix of the General k Class Estimators of the Parameters in Simultaneous Equations" *Econometrica*, 27 575-595
- Nickell, S (1981) "Biases in Dynamic Models with Fixed Effects" *Econometrica*, 49 1417-1426
- NIST/SEMATECH (2013) *e-Handbook of Statistical Methods* <http://www.itl.nist.gov/div898/handbook/>, date.
- Ramsey, J.B (1969) "Regression Specification Error Test" *Journal of the Royal Statistical Society. Series B*, 31 350-371
- Siegel, S. (1994) *Estadística No Paramétrica Aplicada a las Ciencias de la Conducta* Trillas 2º ed.
- Silk, A.G., Urban, G.L (1978) "Evaluation of New Packaged Goods: A Model and Measurement Methodology" *Journal of Marketing Research* 25(4) 171-191
- Snedecor, G.W., Cochran, W. (1989) *Statistical Methods* 8º edition Iowa State University Press
- Tellis, G.J. (1988) "The Price Elasticity of Selective Demand: A Meta-Analysis of Econometric Models for Sales" *Journal of Marketing Research* 25(4) 331-341
- Van Heerde, H.J., Leeflang, P.S.H., Wittink, D.R. (2004) "Decomposing the Sales Promotion Bump with Store Data" *Marketing Science*, 23(3) 317-334

- Vidale, H.L., Wolfe, H.B., (1957) "An operations-research study of sales response to advertising" *Operations Research* 5 370-381
- Wansbeeck, T.J. and Kapteyn, A. (1982) "A simple way to obtain the spectral decomposition of variance component model for balanced data" *Communications in Statistics*, A11 2105-2112
- Wierenga, B. (ed) (2008) *Handbook of Marketing Decisions Models* Springer
- Wittink, D.R., Addona, M.J., Hawhes, W.J., Porter, J.C. (1988) "SCAN\*PRO: The estimation, validation and use of promotional effects based on scanner data". *Internal paper*, Cornell University, Ithaca, NY.
- Wooldridge, J.M (2002) *Econometric analysis of cross section and panel data* Massachusetts Institute of Technology
- Wooldridge, J.M (2005) *Introducción a la econometría: Un enfoque moderno* 2ª edición Thomson 298
- Yustas López, Y (2014) *El producto en la tienda: El panel de detallistas en la investigación de mercados* Universidad Pontificia Comillas